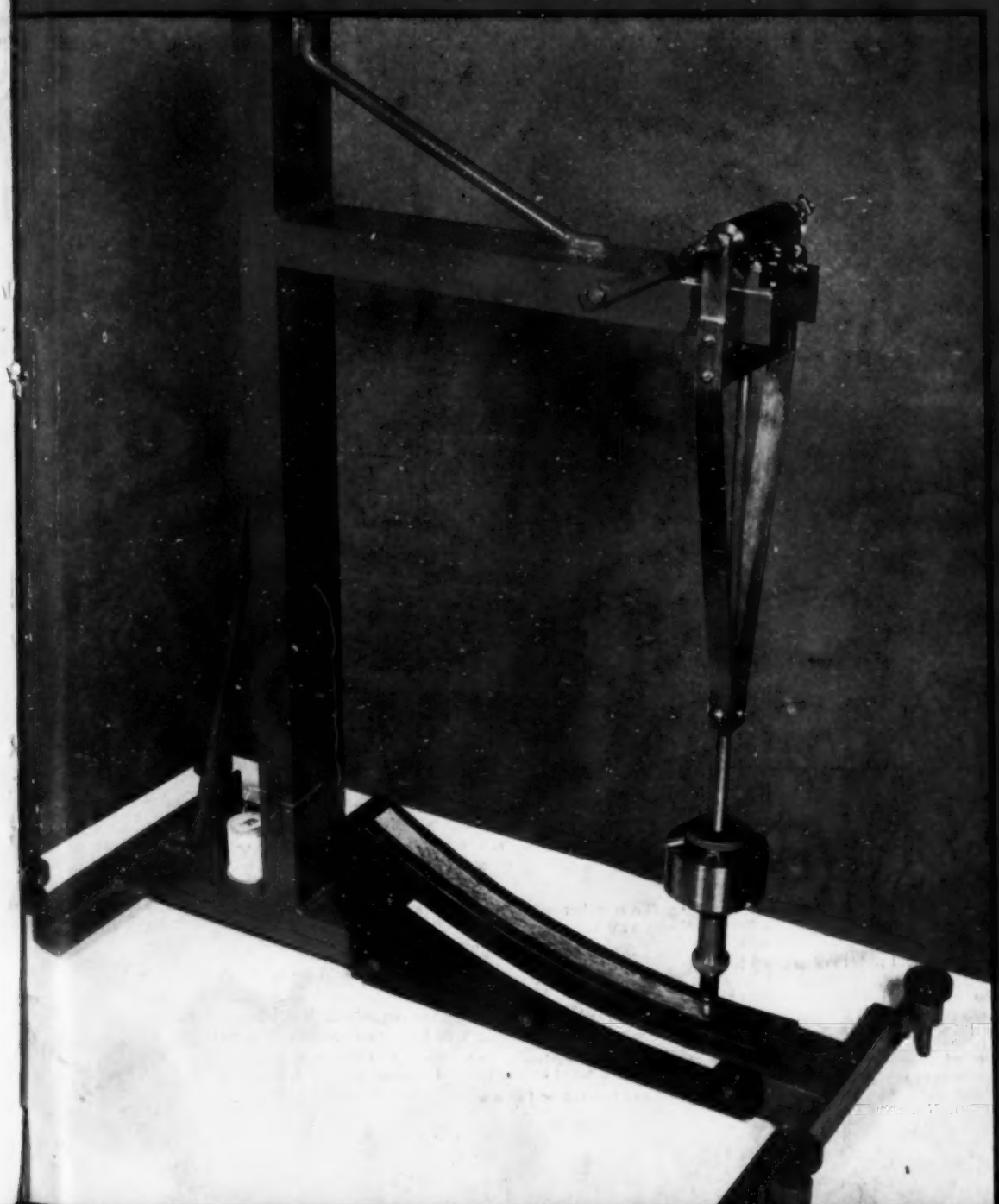
TECHNOLOGY DEPT.

PUBLIC LIBRARY
JUL 10 1941
DETROIT

JOURNAL OF APPLIED PHYSICS



JULY
1941
Vol. 12 No. 7

Physical Aspects of Boundary Lubrication
Otto Beeck

JOURNAL
OF
APPLIED PHYSICS

JULY, 1941 Vol. 12, No. 7

In This Issue

		Page
Radiolocators		511
Physical Aspects of Boundary Lubrication	OTTO BEECK	512
Electron Theory of Thermoelectric Effects	W. V. HOUSTON	519
Frictional Phenomena. I	ANDREW GEMANT	530
Theory of the Plastic Properties of Solids. IV	FREDERICK SEITZ AND T. A. READ	538

Contributed Original Research

Temperature Correction Methods in Calorimetry	ALLEN KING AND HORACE GROVER	557
Kinetics of Degradation and Size Distribution of Long Chain Polymers	ROBERT SIMHA	569
Consolidation Settlement of a Soil with an Impervious Top Surface	M. A. BIOT AND F. M. CLINGAN	578

In Every Issue

Here and There	555
New Book	556
Innovations in Instruments	582
New Instrument Booklets	vi
Index to Advertisers	viii

BOARD OF EDITORS

ELMER HUTCHISSON,
Editor

Associate Editors

FRANCIS BITTER
W. F. BUSSE
ROBLEY D. EVANS
MORRIS MUSKAT
B. E. WARREN
J. B. WHITEHEAD

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

HENRY A. BARTON,
Director

MADELINE M. MITCHELL,
Publications Manager

JOHN T. TATE,
Adviser on Publications

GOVERNING BOARD

PAUL E. KLOPSTEG,
Chairman
GEORGE B. PEGRAM,
Secretary
KARL T. COMPTON
WHEELER P. DAVEY
FLOYD A. FIRESTONE
HARVEY FLETCHER
R. C. GIBBS
ARTHUR C. HARDY
GEORGE R. HARRISON
A. STUART HUNTER
H. R. LILLIE
A. G. WORTHING
W. B. RAYTON
JOHN T. TATE
WALLACE WATERFALL

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. It aims to be of service not only in physical laboratories, but also in the laboratories of industrial, chemical, geological, geophysical, meteorological, radio, and similar concerns. The JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, American Institute of Physics, 175 Fifth Avenue, New York, New York.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 175 Fifth Avenue, New York, N. Y.

SUBSCRIPTION PRICE

	U. S. AND CANADA	ELSEWHERE
To members of scientific societies	$5.00	$5.70
To all others	7.00	7.70

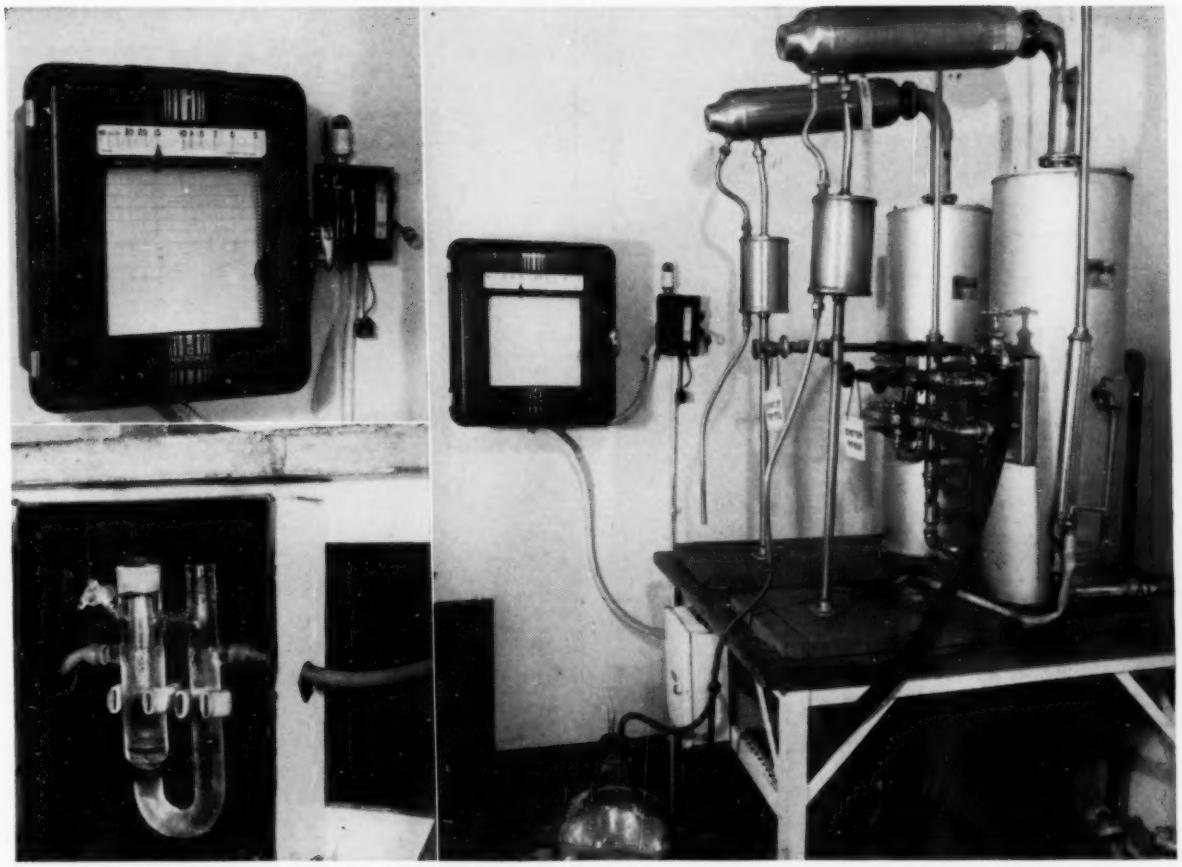
Single copies are $.70 each.

Changes of address and proofs should be addressed to the Publications Manager.

Advertising rates supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1897.
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1941, by the American Institute of Physics.



Automatic equipment providing water for intravenous and other parenteral solutions in Jefferson Hospital, Philadelphia. One of the two stills is for standby service. Micromax recorder and cell, which record purity of distillate from still in use, are shown close-up at left.

Hospital Has Continuous Proof of Water Purity

Among many interesting features of the Jefferson Hospital centralized sterilization room is an automatic device for checking and recording the purity of distilled water.

Water is provided by a ten-gallons-per-hour Barnstead automatic still. The effluent flows through a conductivity cell which is connected to a Micromax strip-chart recorder. This combination measures and records the electrical resistance of the distillate. A resistance of 82,000 ohms per cc. has been adopted by this hospital as satisfactory for parenteral use, and since the conductivity cell used with the equipment has a constant of 0.1, the resistance shown by the recorder becomes 8200 ohms. When the still is started and until the resistance is above this figure, the red light is on. If the effluent should drop again to that figure, the recorder would flash a red light to summon one of the station nurses.

Principal advantages of this method of analysis are that it largely eliminates the human equation, and that it provides a graphic record, always available for consultation. It checks every drop of water instead of occasional grab samples. It wastes no water. Since the purity is recorded, the station supervisor can watch not only for sudden or temporary changes but for gradual ones due to aging of equipment, seasonal differences in raw water, etc.

Catalog EN–95, sent on request, describes the equipment.

LEEDS & NORTHRUP COMPANY, 4978 STENTON AVE., PHILA., PA.

LEEDS & NORTHRUP

MEASURING INSTRUMENTS · TELEMETERS · AUTOMATIC CONTROLS · HEAT-TREATING FURNACES

Jrl Ad E(18)

"IS THE STANDARD CELL AN EPPLEY CELL?"

The far-seeing purchasing agent will find it to his advantage to ask this question before making a decision on the purchase of automatic temperature control apparatus because—

For nearly a quarter of a century the leading American manufacturers of such apparatus have used Eppley Standard Cells in their recorders, pyrometers and other instruments utilizing the potentiometric system of measurement.

Today, more than ever before, these manufacturers realize that the standard cell—the heart of the potentiometric system—must be the finest and most reliable cell obtainable.

The technical men responsible for the performance of the instruments know that the Eppley Standard Cell measures up fully to their standards of excellence and reliability.

EPLAB

THE EPPLEY LABORATORY, INC.
SCIENTIFIC INSTRUMENTS
NEWPORT, R.I.
U.S.A.

INCREASE EFFICIENCY

UNCOATED COATED

SPECTROSCOPES
MICROSCOPES
OPTICAL INSTRUMENTS

By high vacuum coating with suitable fluoride film* to minimize surface reflection loss shown above in half coated aircraft instrument.

NATIONAL RESEARCH CORPORATION
100 BROOKLINE AVE.
BOSTON, MASS.

*U. S. Pat. 2,207,656

GEOPHYSICS

Volume VI, Number 2
(April, 1941)

Application of the Fourier Integral to Some Geophysical Instrument Problems
W. T. Born and J. M. Kendall

The Effect of the Placement of a Seismometer on its Response Characteristics
Harold Washburn and Harold Wiley

The Attenuation Constant of Earth Materials
W. T. Born

The Range of Amplitudes in Seismic Reflection Records *J. M. Kendall*

Average Vertical Velocities from Refraction and Reflection Profiles
W. R. Ransone and Frederick Romberg

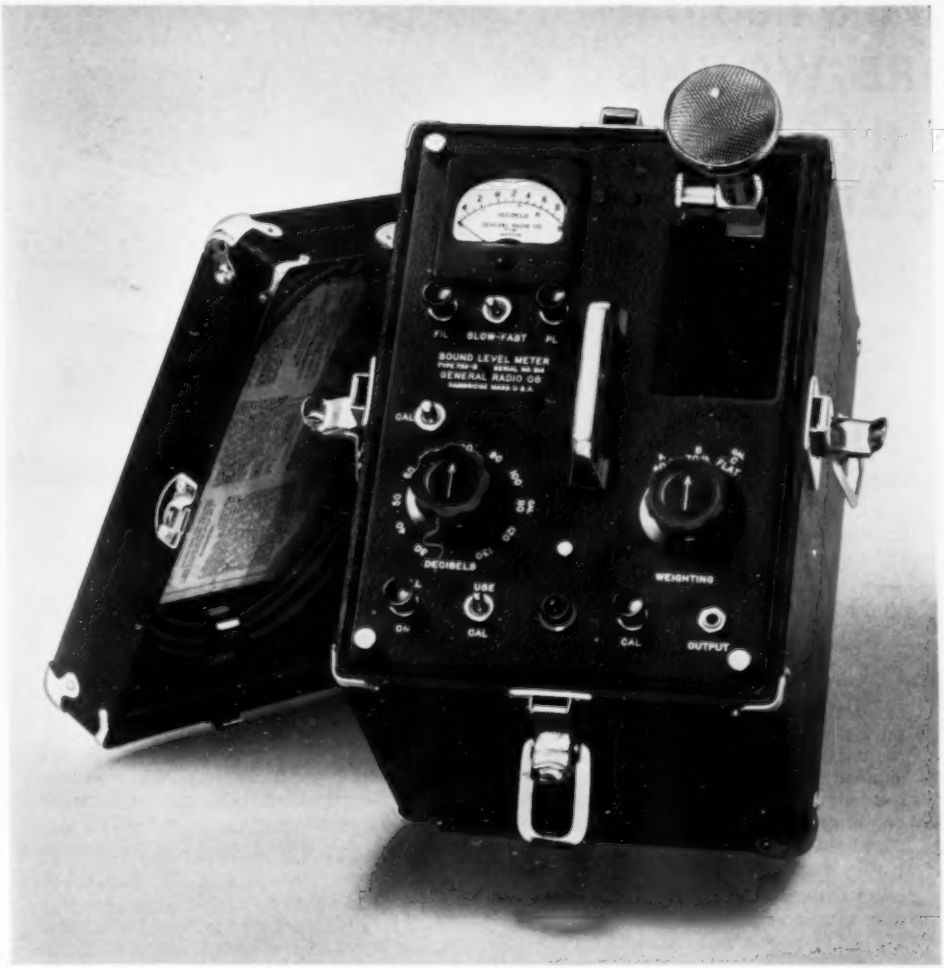
Network Adjustment by Least Squares—Alternative Formulation and Solution by Iteration
M. O. Gibson

The Calculation of Gravity Anomalies due to Bodies of Finite Extent *S. Levine*

An index of past Society publications may be obtained by addressing

THE SOCIETY OF EXPLORATION GEOPHYSICISTS
P. O. Box 777 **Austin, Texas**

Simplified Sound Measurements



With this Accurate, Highly Stable Meter

WITH THE IMPROVED General Radio Sound-Level Meter the intensity of sound can now be measured conveniently, quickly and accurately by anyone. This instrument is widely used throughout industry for all types of noise measurements and studies. It is completely self-contained and portable and will measure the intensity of sound from a whisper to a steam whistle.

FEATURES:

1 EXCEPTIONALLY WIDE RANGE—24 to 140 db

2 SIMPLIFIED CONTROLS—complete db range covered by a single control knob

3 IMPROVED MICROPHONE—special type of great sensitivity, highly rugged, unaffected by ordinary changes in temperature and humidity

4 PLUG-IN MICROPHONE—normally used directly on folding socket on instrument but can be detached and used with cable and tripod if desired

5 SLOW-FAST METER—two speeds—control switch selects either; "fast" position, meter movement corresponds to A. S. A. standards; "slow" position, meter heavily damped to measure average level fluctuations

6 INTERNAL CALIBRATION SYSTEM—calibration can be checked quickly from built-in calibrator

7 THREE WEIGHTING NETWORKS—low level, high level and substantially flat over-all response

8 NO COILS—instrument contains no coils or inductances and accordingly reads accurately in the presence of any ordinary magnetic field

TYPE 759-B SOUND-LEVEL METER $195.00

• Write for Bulletin 718 for Information

GENERAL RADIO COMPANY, Cambridge, Massachusetts

PRECISION WIRE WOUND RESISTORS

The Answer to Many Bridge Circuit Resistance Requirements

Never has there been such a concerted development of electrical equipment requiring resistors for bridge and similar circuits where the call is for greater accuracy and lower temperature co-efficient than heretofore available in production quantities.

IRC Precision Wire Wound Resistors were designed for such purposes. They are made to a standard tolerance of ±1% or to as low as 1/10 of 1% on special order. They have an extremely low temperature co-efficient for stable operation over a temperature range from -30° C. to 100° C. They are impregnated against humidity and other atmospheric conditions. They are available in the most complete line of inductive and non-inductive types, shapes, sizes and terminals on the market today.

Thousands of these units used in a wide range of equipment from voltmeter multipliers for sea coast duty to geophysical instruments used in tropical climates prove their dependability under the most exacting conditions.

Whatever your resistance requirements, it pays to ask IRC. IRC engineers will welcome the opportunity to cooperate. They are backed by the world's largest line of fixed and variable resistors. Their recommendations are based on many years of specialized experience in dealing with practically every type of resistance problem.

Write for IRC Resistance Engineering Data Bulletin No. IV including both IRC Precision and Cement Coated Power Wire Wound Resistors.

MAXIMUM PROTECTION • CLOSE TOLERANCES • SMALLER SIZES

This is an actual size view of an IRC Type BW ½-watt Insulated Wire Wound Resistor — a compact, well-protected unit that can often be used in exacting applications where space is at a premium. Other Type BW sizes are 1-watt and 2-watts. Tolerances available to as close as 2%. Some sizes and ranges can be supplied to even closer tolerances and in matched pairs to 1%.

IRC
REG. U.S. PAT. OFF.

INTERNATIONAL RESISTANCE CO. 419 NORTH BROAD ST. PHILADELPHIA, PA.

Journal
of
Applied Physics

Volume 12 July, 1941 Number 7

RADIOLOCATORS

IN a radio appeal* on June 18 for volunteers for civilian technical workers Lord Beaverbrook, British Minister of State, released some of the secrets of the "radiolocator," a device which, more than anything else, enabled Great Britain to ward off the ferocious Nazi air attack last fall. According to statements released to the press, the radiolocator sends out radio-waves which are essentially unaffected by fog, clouds or darkness. Any object in the path of the waves, such as an enemy plane, reflects back a signal which locates the plane and gives instant warning of attack.

A statement by the British Air Ministry includes the following: "The development of the radiolocator has created a new military science which came into being without the world being aware of its existence and which has already exerted a profound influence on air, military and naval strategy. Today production lines in the plants of radio manufacturers throughout Great Britain are turning out radiolocators in ever-increasing quantities, and their installation throughout the country has made it imperative to train more men to service and maintain them, although thousands are already engaged in this work. . . . They are not simple pieces of apparatus. Before he is competent to maintain one a man needs to have a sound working knowledge of radio to begin with, and then he must have at least two months of specialized training. Thousands of men with a knowledge of radio are required to look after the radiolocators now being produced. The use of the radiolocator in this system has been so successful that independent observers who have seen it in operation have expressed the opinion that, but for it, the first Battle of Britain might have been lost many months ago. But the first battle was not lost, and the system of radiolocation is now being widely extended. Radio men, accepted for this work, receive a course of technical instruction lasting approximately eight weeks. However much they may already know about radio, this course and the experience that will follow will extend their knowledge of new developments in radio far beyond anything yet imagined. . . ."

The British themselves need immediately tens of thousands of operators and service men for radiolocators now coming off the production lines. What our own needs will be it is difficult to estimate, but we can be sure they will be enormous. There seems to be no question but that we must train many more workers in electronics and radio. Funds for the cost of instruction will undoubtedly be available through legislation sponsored by the U. S. Office of Education and now before Congress. Unfortunately, as yet no federal provision is being made for scholarships or fellowships to attract some of the most outstanding graduates to advanced study in this field.

* For more details of this appeal see the *New York Herald Tribune* or *The New York Times* for June 18 and 19, 1941.

For August

Special Issue on Photoelasticity

including

Three-Dimensional Photoelasticity R. WELLER

Stress and Flow Patterns MIKLÓS HETÉNYI

Numerical Solution of Laplace's Equation M. M. FROCHT AND M. M. LEVEN

Classification of Isotropic Points MICHAEL A. SADOWSKY

Stresses in Gear Teeth THOMAS J. DOLAN

and several other short papers and abstracts.

The Physical Aspects of Boundary Lubrication

BY OTTO BEECK

Shell Development Company, Emeryville, California

What is Boundary Lubrication?

IF two lubricated surfaces are sliding over each other under conditions of high pressure and low speed, no truly fluid film of the lubricant can exist between the sliding surfaces, and the term "boundary lubrication" has been widely adopted for this condition. Boundary lubrication or non-hydrodynamic lubrication is therefore readily distinguished from "fluid film" or "hydrodynamic" lubrication, which takes place under maintenance of an oil film of sufficient thickness so that hydrodynamic properties such as viscosity and the relative motion of the two surfaces are the sole characterizations necessary. Under boundary conditions, phenomena present themselves which must be associated with physical characteristics of the lubricant other than viscosity. They are, in fact, associated with chemical as well as physical characteristics of both the lubricant and the surface, which makes boundary lubrication a very elusive and complicated scientific problem. Lack of fundamental knowledge in this field together with a strong realization of its practical importance led some time ago to the term "oiliness," which simply indicates the lubricating value of an oil under boundary condition; this term is still widely used. Its definition, as adopted later by the Society of Automotive Engineers, is as follows:

"Oiliness is a term signifying differences in friction greater than can be accounted for on the basis of viscosity when comparing different lubricants under identical test conditions."

The state of boundary lubrication should, however, be defined much more stringently. In the author's opinion, the best criterion for this state is that the coefficient of friction is independent of viscosity and of the sliding velocity. Pure boundary lubrication thus defined is extremely rare in practice because quasi-hydrodynamic effects of the lubricant are difficult to eliminate but we shall see that it can be obtained in the laboratory.

Considerable progress has been made within the last 10 years in the understanding of the phenomena of boundary lubrication and a good deal has been contributed by physicists and physical chemists. It is primarily their efforts which will be discussed in this summary without any claim of completeness.

The Area of Contact Between Stationary and Sliding Surfaces

It is obviously quite important that before speaking of friction between sliding surfaces we should have some idea of the actual area of contact of such surfaces under various loads, since it is this area which determines the actual pressure exerted on a protective boundary film between the surfaces. Bowden and Tabor[1] (1939) have made an extensive study of this subject. They found that the real area of contact of dry surfaces is not greatly affected by size, shape and even degree of roughness of the surface and that it depends mainly on the pressure. This is true also for surfaces which are very carefully ground and polished. The real area of contact was estimated by electrical and by visual means and the following results were obtained for steel flats of 21 cm^2.

LOAD, KG	FRACTION OF MACROSCOPIC AREA IN CONTACT
300	1/130
100	1/700
20	1/10,000
3	1/170,000

This means that even for lightly loaded surfaces the local pressure may be sufficiently great to cause plastic deformation of the steel. A careful investigation of this point through measurements of the conductivity of crossed steel or copper cylinders leaves no doubt that plastic deformation takes place at the local points of contact and that the real area of contact between flat surfaces is proportional to the normal load, whereas on the basis of elastic deformation the real area of contact should vary as the two-thirds

power of the normal load. This, at once, gives an explanation of the old law of Amontons (1699) which states that the tangential frictional force is directly proportional to the normal load, the mysterious meaning of which had only recently been discussed by Adam[2] (1938). It also explains the well-known other observation of Amontons that the friction is independent of the macroscopic area of the surface.

When the real area of contact of two sliding surfaces is measured by the electrical conductivity methods, rapid fluctuations occur as would be expected from an intermittent making and breaking of the contacts. However, *the type and the magnitude of the fluctuations remain essentially the same when the surfaces are lubricated under boundary conditions with mineral oils or other lubricants.*

The Surface Temperature of Sliding Metals

It is clear from the preceding paragraph that when two surfaces slide over each other, the actual pressure on the real area of contact may

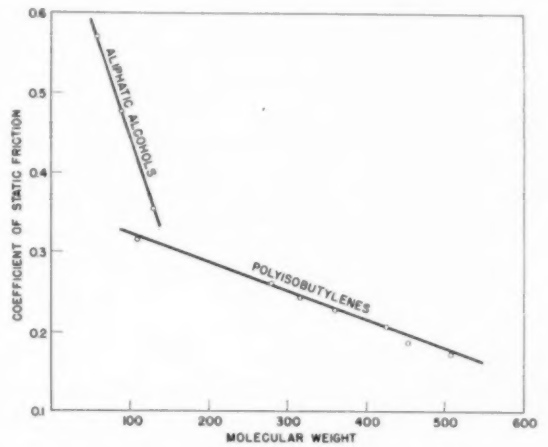


FIG. 1. Coefficients of static friction showing linear relationship with molecular weight for homologous groups.

exceed several thousand kg/cm² according to the pressure at which plastic deformation occurs in individual cases. As most of the work done against the frictional force is liberated as heat in the very small area of real contact, a measurement of the temperature at the points of contact should give information about the actual temperature of the surface atoms of the rubbing metals. Bowden and Ridler[3] (1936) have calculated and measured such surface temperatures. Calcuation for the simple example of a constantan cylinder 1 mm in diameter loaded with 100 g and sliding over a mild steel surface with a velocity of 100 cm per second and having a measured value of the coefficient of friction of 0.23 gives the following results.

Fraction of Surface Area in Rubbing Contact	Calculated Surface Temperature
All of the surface	75°C
1/10	414
1/100	2372

The measurements were done by sliding two different metals over each other and measuring the electromotive force set up by the resulting thermocouple.* The temperatures reached depend upon load, speed, coefficient of friction and thermal conductivity of the materials used and are in good agreement with the theory. With metals of low melting point, the surface temperatures reached correspond to the melting point. With less fusible metals surface temperatures exceeding 1000°C were observed. *Lubricating the surfaces under boundary conditions did not prevent the surfaces from reaching temperatures of 600°C,* and it is evident that breakdown or decomposition of the boundary film must have occurred.

Friction Under Boundary Conditions

The most direct experimental evidence that friction of lubricated surfaces under boundary conditions is intimately connected with the physico-chemical properties of the molecules themselves was obtained by Sir William Hardy and his co-workers (1922–25) in their classical experiments on static friction using as lubricants the homologous series of paraffins, alcohols and fatty acids. The relationship obtained was unexpectedly simple and the friction was found to be a function of separate contributions by the solid surfaces, the chemical series, and the number of carbon atoms in the chain. The results were interpreted on the assumption that friction took place between monomolecular films of lubricant adsorbed on the solid surfaces. Figure 1 shows some data obtained by Dr. Givens

* Blok (1937) (see reference 4) has used the thermocouple principle to measure "flash" temperatures on gear teeth during meshing and has estimated the highest temperatures to be of the order of 200–250°C. He also has made a general theoretical study of the temperature rise at surfaces of actual contact under boundary conditions.

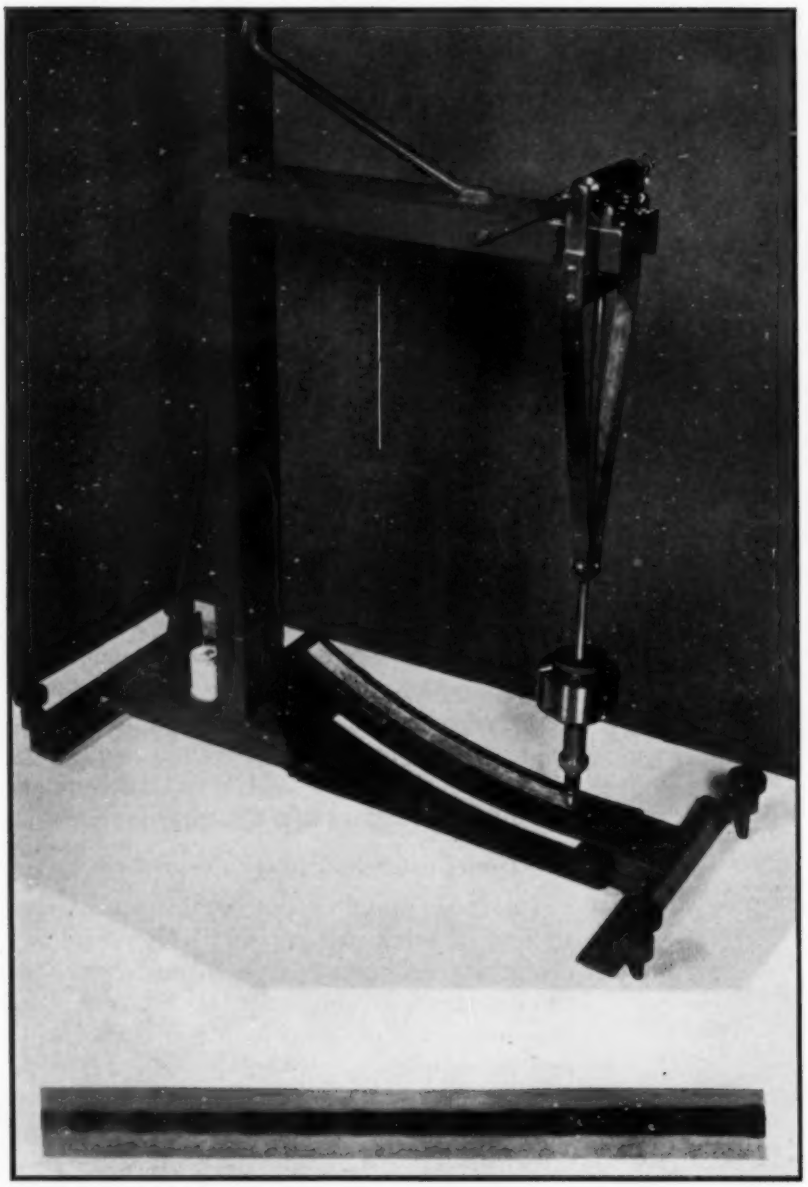


FIG. 2. Friction pendulum with photographic record of decrement.

in these laboratories. Hardy's experiments are so convincing that for many years static experiments have been employed to measure "oiliness" or the properties of lubricant under boundary condition. The practical value of these data is very doubtful, however, as they do not reveal any properties of adsorbed films under sliding conditions. In fact, we have seen in the preceding paragraphs that under conditions of continuous sliding very high temperatures are produced at the local points of contact, and it seems quite impossible that under these severe conditions any monomolecular film could stand up unless it were self-repairing. Langmuir[10] (1934) has described some striking experiments to illustrate this point. When a clean spherical glass rider is placed on a clean flat glass slide and the slide is tilted, the rider will not slide down unless the angle is at least 60°. When a monomolecular film of fatty acid is placed on the rider, slipping will occur at 40° but the rider will only slip a few millimeters after which the angle must again be raised to 60°. When, however, the monomolecular film is applied to the slide, an angle of only 6° to 8° is necessary to cause the rider to slip off the slide irrespective of whether a film also had been applied to the rider. This is of course due to the fact that the rider moves continuously to fresh parts of the surface, whereas in the first experiment, the same point of contact is maintained with respect to the rider and the film applied to the rider wears out almost immediately. In another experiment, the rider moved over the slide in reciprocating motion on the same track. If a monolayer of fatty acid was placed on the slide, it took 700 strokes to increase the coefficient of friction from 0.11 to 0.13 and an additional 200 strokes to raise it to 0.24. With a film 7 layers thick the coefficient of friction did not increase perceptibly even after 4000 strokes. Similar experiments were carried out by Dr. Givens and the author using a friction pendulum pictured in Fig. 2 with either a journal bearing

on four balls or a four-ball bearing (Fig. 3).† The balls and the journal were covered with varying numbers of monolayers of fatty acids or related compounds like octadecyl stearate or tristearin by the Langmuir-Blodgett method. Coefficients of friction could be obtained by measuring the decrement of the pendulum using the photographic record shown in Fig. 2. In agreement with Langmuir it was found that no less than seven layers would give a small and constant decrement. Seven layers, in fact, seemed to wear indefinitely.‡

It was already pointed out that with flooded bearings of any type true boundary lubrication may be difficult to obtain, and if obtainable the coefficient of boundary lubrication should be independent of the velocity of the rubbing surfaces. Experiments with reciprocating motion are not very suitable to investigate this point. The first clean-cut case of boundary lubrication was obtained by Beare and Bowden[7] using a flat flywheel with a spherical rider. The flywheel was set in motion and its deceleration was observed by plotting the angular velocity against time. With highly polished steel surfaces and fluid lubricants such as octyl alcohol, a straight line was obtained. Viscous and oily lubricants, however, showed always a dependence of their coefficients of friction upon velocity, even at very low velocities. We have shown at Emeryville (Beeck, Givens and Smith[8]), that even with viscous and oily lubricants in the flooded state true boundary conditions can be obtained. The apparatus used was the Boerlage four-ball top shown in Fig. 4. This apparatus in which a single ball rotates under variable loads on a support formed by three similar balls clamped together in an oil cup is particularly adaptable to measurements of coefficients of friction at low sliding velocities of 0 to 1 cm/sec. and high pressures of a few thousand kg/cm^2 or higher (as calculated from the elastic properties of the steel balls).

A careful analysis of the motion of the rotor was made by means of a photoelectric device in conjunction with a revolving drum recorder. Instantaneous values for the deceleration were obtained by graphical differentiation. The results were finally expressed in plotting coefficients of friction against velocity. Figure 5 shows some typical results. At certain critical and easily reproducible velocities, the curves bend over sharply and the coefficient of friction becomes constant. Ricinoleic acid, the main constituent of castor oil, does not show this effect. Neutral or white oils likewise do not show the effect at low velocities. It is imparted to these oils by the presence of small amounts of such fatty acids or other compounds that are readily adsorbed in form of well oriented surface films as shown by electron diffraction experiments.

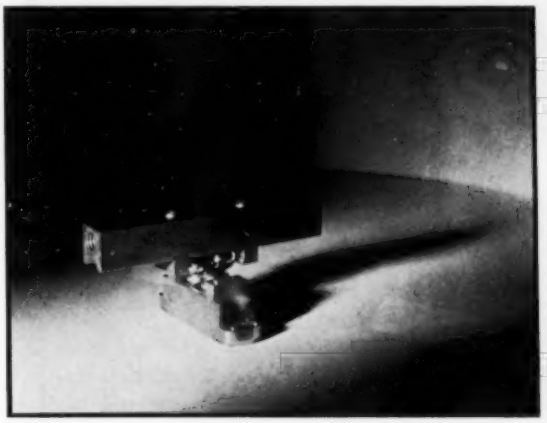


FIG. 3. Four-ball bearing for friction pendulum.

Since built-up films of fatty acids of one or a few molecular layers will not produce the effect, it has been concluded that the adsorbed layer of long chain polar molecules is able to induce a *wedging effect* which allows an early transition from the boundary state into a *quasi-hydrodynamic* state in which lower coefficients of friction prevail. Rough surfaces do not show the wedging effect as would be expected for two reasons. The interlocking of the asperities of such surfaces and the extreme pressure and temperature conditions thus produced upon sliding would make the action of an adsorbed layer impossible, and the fact that the degree of orientation of an adsorbed layer depends greatly on the degree of perfection of the surface would make the induction of the wedging effect less probable.

† A friction pendulum was described by Kyropoulos and Shobert (reference 5). The four-ball arrangement and the photographic recording method are the author's.

‡ See also recent experiments with built-up films by Bowden and Leben, reference 6.

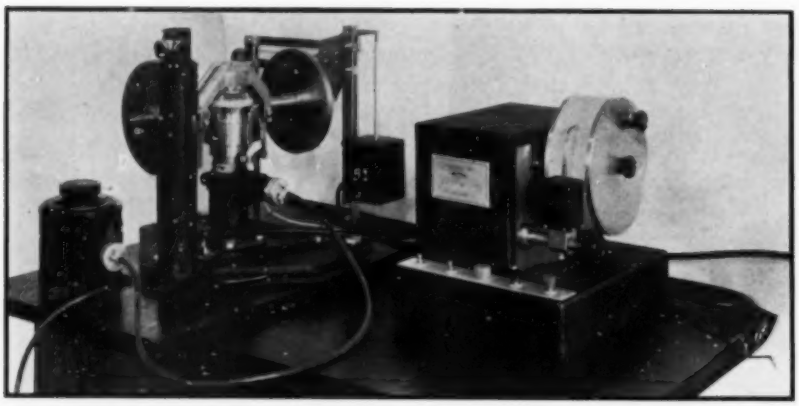


FIG. 4(a). Four-ball top with photoelectric chronograph.

We have seen in an earlier paragraph that Bowden and collaborators have found rapid fluctuation of the surface temperature under boundary lubrication conditions. This makes it quite probable that with an apparatus capable of measuring frictional fluctuations, correlations between frictional and temperature fluctuations may be found. It is essential for such an apparatus that the friction measuring device should have a high natural frequency and that the system which applies the load should have a still higher natural frequency. Bowden and Leben[6] constructed such an apparatus and found the frictional fluctuation which they called "stick-slip." The exact behavior depends on the combination of metals, particularly on their melting points and, of course, on the lubricant if the surfaces are lubricated. In case of dissimilar metals the fluctuatious are quite regular. The surface of the rider sticks to the moving bottom plate (it is, in fact, welded to it) and moves with it until, as a result of the increasing force which wishes to keep the rider in its original position, a sudden and very rapid slip occurs. Simultaneously with this slip there is a sudden temperature flash which produces the next welding, and so on. Surfaces lubricated with straight mineral oils behave in the same manner. However, long chain fatty acids, for instance, may prevent the stick-slip and allow continuous sliding to take place. The experiments of Bowden and co-workers were carried out at very low sliding speeds, and the slider had a low mass. It would be wrong to generalize that all boundary friction occurs with the type of stick-slip just described. Blok[4] (1940) has rightly suggested that stick-slip is a regime which may or may not be entered depending on dynamical characteristics of the sliding machine parts and the relation of static and kinetic coefficients of friction. In the last analysis it is, of course, true that the elastic properties of the restraining system may be that of the surface irregularities themselves in which case stick-slip of distances of only a few atoms and periods of 10^{-4} to 10^{-5} second may still occur.

Boundary Lubrication and Wear

Although sliding friction and mechanical wear are logically closely connected, no significant general correlation is found in practice between the coefficients of friction and wear with various oils with and without addition agents. Under moderate conditions, the coefficient of boundary friction varies but little with pressure, temperature, and material of both lubricant and surfaces. The reason for this lies again in the fact that wear

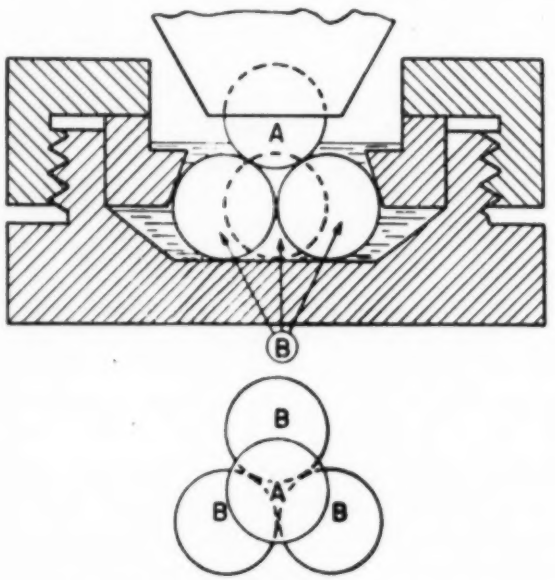


FIG. 4(b). Cross section of four-ball bearing.

takes place momentarily at isolated spots which are small compared to the whole surface, and although these isolated spots must exhibit a higher coefficient of friction at the moment of abrasion, the resulting change for the whole lubricated system is too small to affect the over-all coefficient of friction unless refinements of the type of Bowden's "stick-slip" method are used.

Wear is, however, an accumulative effect and therefore its measurement is always possible even if the rate of wear is very low. Furthermore, the engineer is seldom interested in the relatively small variations in power loss due to variations in the coefficient of friction in machine parts which partly or accidentally have to rely on boundary lubrication, but he is interested in wear. We have, therefore, attempted at Emeryville to investigate wear (Beeck, Givens, Williams[9] (1940)) under conditions which would enable us to link friction and wear measurements more closely than has heretofore been possible. A motor-driven modification of the four-ball top was used as shown in Fig. 6. As wear was unmeasurably low at the low speeds at which the wedging effect was observed, speeds of several hundred revolutions per minute had to be used. A successful innovation was the use of steel balls which were electroplated with a very thin and uniform copper plate, whose wear characteristics were found to be similar to those of iron, and which were simply run until the track on the upper rotating ball was worn through to the steel. This time could be easily measured, and the experiments were found to be reproducible within 1 percent. Under these conditions long chain polar compounds contributed very little to wear prevention, but it was found that other known wear prevention agents, like tricresyl phosphate, showed a wear reduction factor of about five. A careful study, especially by Dr. Givens, of the chemistry of the whole metal-lubricant system led to a complete understanding of the mechanism of the action of tricresyl phosphate or similarly acting compounds. Their action is essentially that of chemically polishing the surface and thereby distributing the load and lowering the surface temperature. The formation of low melting metal phosphide alloys plays a major role in the mechanism of this action. Here we cannot fail to be reminded of the interesting work of Bowden and Hughes (1937) who have investigated the mechanism of polishing, surface flow, and the formation of the Beilby layer. They find that the process of polishing is mainly influenced by the relative melting point of the polisher and the solid and that the relative hardness is comparatively

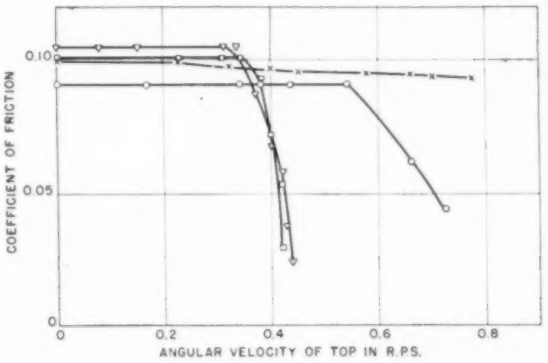


FIG. 5. Coefficients of friction obtained with four-ball top. ○ castor oil; × ricinoleic acid; ▽ fatty acid fraction from castor oil; ⊡ 1 percent stearic acid in ricinoleic acid.

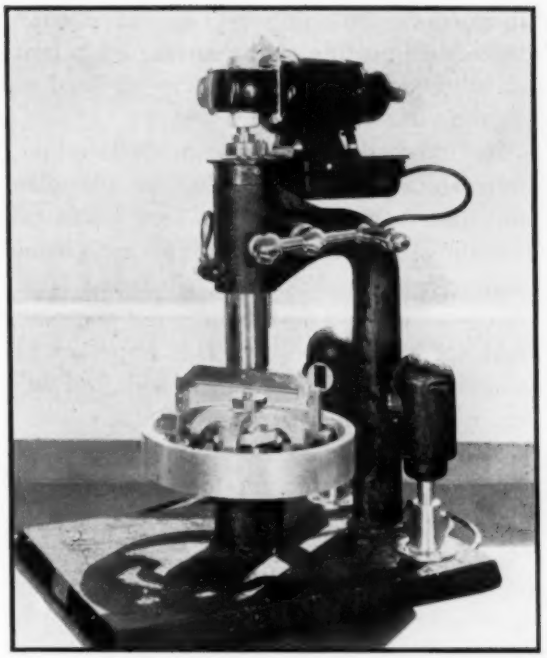


FIG. 6. Four-ball wear apparatus.

unimportant. Polishing is brought about by intense local heating of the surface irregularities and subsequent smearing of the softened solid over the surface.

After having gained some understanding of the chemical polishing action of phosphorus compounds and the action of long chain polar compounds, it seems suggestive that chemical polishing agents should create just that type of surface on which long chain polar compounds would act most favorably. This hope has been justified beyond expectation. A white oil whose wear reduction factor was raised to five by addition of 1.5 percent tricresyl phosphate showed a wear reduction factor 10 when, in addition to the phosphorus compounds, 1 percent of oleic acid was added. Factors as high as 17 have been found with other combinations. This self-consistency of results is most gratifying in a field where in practice the multiplicity of conditions may easily obscure the application of fundamental knowledge.

Surface Finish

It would indeed be difficult not to say in this connection a word about surface finish, although the field is so large that it can only be touched upon. It is sufficient to say, however, that engineers have been well aware of the fact that surface finish has an important bearing on wear; and the general tendency of the last 10 years has been towards better surface finishes of lubricated rubbing surfaces, even if such surfaces are operating most of the time under hydrodynamic conditions. The importance of this field was well demonstrated at a special Summer Conference on Friction and Surface Finish at M. I. T. June 6–7, 1940 (see proceedings of this meeting). It is particularly significant in the light of the work on boundary lubrication that the mechanical approach from the standpoint of surface finish and the physico-chemical approach from the standpoint of the addition agents have both led to the realization that mechanical wear is mainly a function of the quality and maintenance of surface finish.

Bibliography

1. F. P. Bowden and D. Tabor, Proc. Roy. Soc. **169**, 391 (1939).
2. N. K. Adam, *The Physics and Chemistry of Surfaces* (Oxford University Press, 1938), p. 232.
3. F. P. Bowden and K. E. W. Ridler, Proc. Roy. Soc., **154**, 640 (1936).
4. H. Blok, "General discussion on lubrication," Inst. Mech. Eng., **2**, 14, 222 (1937); S. A. E. Trans. **35**, 54 (1940).
5. S. Kyropoulos and E. I. Shobert, Rev. Sci. Inst. **8**, 151 (1937).
6. F. P. Bowden and L. Leben, Proc. Roy. Soc. **169**, 371 (1939); Phil. Trans. Roy. Soc. **239**, 1 (1940).
7. W. G. Beare and F. P. Bowden, Phil. Trans. Roy. Soc. **234**, 329 (1935).
8. O. Beeck, J. W. Givens and A. E. Smith, Proc. Roy. Soc. **A177**, 90 (1940).
9. O. Beeck, J. W. Givens and E. C. Williams, Proc. Roy. Soc. **A177**, 103 (1940).
10. I. Langmuir, J. Frank. Inst. **218**, 143 (1934).

Electron Theory of Thermoelectric Effects

By W. V. Houston

California Institute of Technology, Pasadena, California

THE development of the theory of solids during the past decade has made possible an understanding, at least in a qualitative way, of the thermoelectric properties of metals.[1] It is now possible to describe, in physical terms, the behavior of the electrons that leads to the different signs and magnitudes of the Peltier effect, the Thomson effect, and the thermoelectromotive force. At the same time it is possible to illustrate the electronic aspect of the two types of electromotive force distinguished by Bridgman,[2] the driving electromotive force and the working electromotive force. Since electromotive force itself is a concept that is frequently misunderstood, this illustration may be of value.

Definitions of the Thermoelectric Coefficients

The thermoelectric effects that we shall consider are the thermoelectromotive force, the Peltier effect, and the Thomson effect. These may be defined in connection with the thermoelectric circuit shown in Fig. 1. In this figure, A and B represent two different metals connected together at the ends. The regions around the junctions are maintained at temperatures T_1 and T_2, respectively. Let T_2 be greater than T_1. The conductor B is cut so that we may consider the open circuit situation; but we may also imagine the ends connected by a resistance, or other energy-consuming device, through which a current can flow. It will be understood, however, that the ends of the conductor on either side of the break, and any apparatus connected between them, are maintained at the same temperature.

When the circuit is open there will be a difference of potential across the cut in conductor B. This potential difference is a measure of the thermoelectromotive force in the same way that the difference of potential between the poles of a battery on open circuit is a measure of its electromotive force. This electromotive force will be designated by $\mathcal{E}_{AB}(T_1, T_2)$ to indicate that it depends upon the two temperatures and also upon the two materials. The order of the subscripts indicates that $\mathcal{E}_{AB}(T_1, T_2)$ is positive if a current tends to flow from A to B at the warmer junction. It has been emphasized by Bridgman that a thermoelectromotive force will exist even if the two conductors are merely strips cut in different directions from the same anisotropic crystal. This fact imposes some important restrictions upon the theory.

If the two parts of B are connected by a resistance, a current will flow around the circuit in consequence of the thermoelectromotive force. This flow of current will be accompanied by an

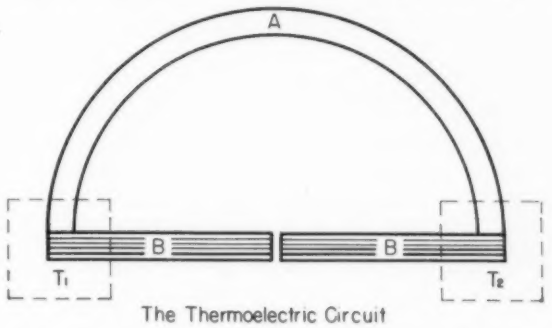


FIG. 1. The thermoelectric circuit. A and B are two different metals whose junctions are kept at the temperatures T_1 and T_2.

absorption of heat at the warmer junction and a liberation of heat at the cooler, superposed on the ordinary liberation of the Joule heat. The rate of liberation or absorption of this thermoelectric heat is proportional to the current. It also changes sign with the current, and if a current is forced around the circuit in a direction opposite to the direction of the thermoelectromotive force, the heat will be absorbed at the cooler junction and emitted at the warmer. This is a process by which the cooler junction is cooled and the warmer junction is warmed by the aid of an external agent. The Peltier coefficient of a junction, $P_{AB}(T)$, depends upon the two metals involved and upon the temperature of the junction. It represents the amount of heat that must be supplied to the junction to maintain its temperature when unit quantity of electricity flows from metal A to metal B.

In addition to the heat absorbed at the junctions, there is heat absorption, positive or negative, along the lengths of the conductors where there exists a temperature gradient. This also contributes to the energy necessary to keep the current flowing, and is known as the Thomson effect. The Thomson coefficient $\sigma_A(T)$ represents

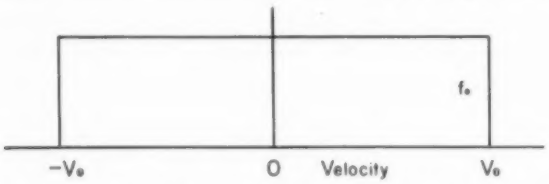


FIG. 2. The normal distribution of electrons in the absence of an electric field and with a uniform temperature.

the amount of heat absorbed by a conductor of metal A at the temperature T when unit quantity of electricity flows in the direction of increasing temperature, and a unit temperature difference exists between the ends of the portion considered. Both the Thomson heat and the Peltier heat are in addition to the Joule heat and must be distinguished from it. They are also in addition to the thermally conducted heat that will always appear in the presence of a temperature gradient, and they are distinguished from it by being proportional to the current. In the theoretical analysis, as well as in experimental work, the thermally conducted heat must be evaluated and subtracted from the total heat energy.

It is these three phenomena, the thermal electromotive force, the Peltier effect, and the Thomson effect, that we wish to understand in terms of the behavior of the conduction electrons. In studying them we must also treat the electrical conductivity, the contact potential, and the thermal conductivity.

Thermal conductivity is ordinarily divided into two parts, one part due to the electrons and one part due solely to the ionic interaction. Only the electronic conductivity of heat will be of importance here, because only it is associated with the electric current.

Statistical Distribution of Electrons in a Metal

The electrons in a metal must be treated statistically, since only their temperature is given. The electronic situation is described by giving the average number of electrons per unit volume in the various states. These states may be specified in various ways, but the simplest way is to regard them as very similar to the energy characteristic states of completely free electrons. The states can then be specified by three quantum numbers, k_x, k_y, k_z that are related to the velocities by the equations

$$v_x = hk_x/m^*, \quad v_y = hk_y/m^*, \quad v_z = hk_z/m^*.$$

The quantity m^* is an effective electron mass. If the electrons are entirely free $m^* = m$, but the effect of the nuclear binding can be roughly described by assigning to m^* values that are somewhat greater or less than m, and that depend on (k_x, k_y, k_z). The kinetic energy ϵ of an electron in such a state is $(h^2/2m^*)(k_x^2+k_y^2+k_z^2)$, and the permitted values of k_x are integral multiples of $2\pi/L$, where L is the length of the piece of metal in the x direction. An electron in such a state has a definite kinetic energy and a definite velocity. It is also assigned a definite potential energy that is taken as constant throughout a single piece of metal. The specification of the statistical distribution is then the specification of the number of electrons having certain velocity components, and this carries with it the specification of the number with any given energy.

When there is no temperature gradient and no electric field, the electrons are distributed over the various states so that the average number in each state is

$$n = 2/\{\exp[(\epsilon-\epsilon')/kT]-1\}, \qquad (1)$$

where k is Boltzmann's constant and T is the absolute temperature. ϵ' is a constant energy that depends upon the temperature and upon the number of electrons per unit volume. For free electrons ϵ' decreases when the temperature is increased, and this may be expected to be true of all good conductors, although the *rate* of decrease will depend upon the nature of the material.

In Fig. 2 the value of n is plotted as a function of k_x, when k_y and k_z are zero. This may be regarded as the plot of n along this particular diameter of a solid figure in the k_x, k_y, k_z space. Each point in such a space represents an electron state, and according to the Pauli exclusion

principle, not more than two electrons (one with each of the two possible directions of spin) can occupy such a state. In assigning the electrons, they are put into the states of lowest possible energy until they are all located. This fills up all the available states below a certain maximum kinetic energy. For perfectly free electrons, the states having the same energy are located on the surface of a sphere, and for the absolute zero of temperature all of the states within the sphere of kinetic energy ϵ' are filled, and all those of greater energy are empty. At higher temperatures the boundary between the occupied and the unoccupied states is less sharply defined, and the function $n(\epsilon)$ falls less sharply to zero. But even at room temperature, the energy spread of this surface region is only of the order of $\frac{1}{2}$ percent of the energy ϵ', and is entirely too small to be shown on the scale of Fig. 2.

In Fig. 3 the portion of Fig. 2 in the neighborhood of $\epsilon=\epsilon'$ is very much enlarged in order to show the way in which n decreases from its maximum value to zero. The value of n as a function of k_x, rather than as a function of ϵ, is plotted, although over the very short region shown, ϵ and k_x are proportional to each other.

Contact Potential

One of the things that it is important to understand in the discussion of the thermoelectric effects, is the relationship between the electron distributions in two different metals, when these metals are brought into contact.

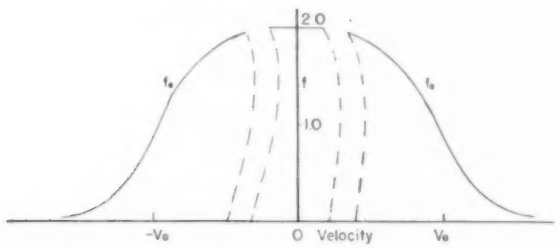


FIG. 3. Enlargement of the electron distribution in the neighborhood of $\epsilon=\epsilon'$.

This involves a study of the contact potential difference and the understanding of why the contact potential difference is not directly connected with the Peltier effect. In fact a good deal of the study of the theory of the thermoelectric phenomena consists in understanding why certain things do *not* produce the effects observed.

Figure 4 illustrates the relationships between two different metals when they are brought into

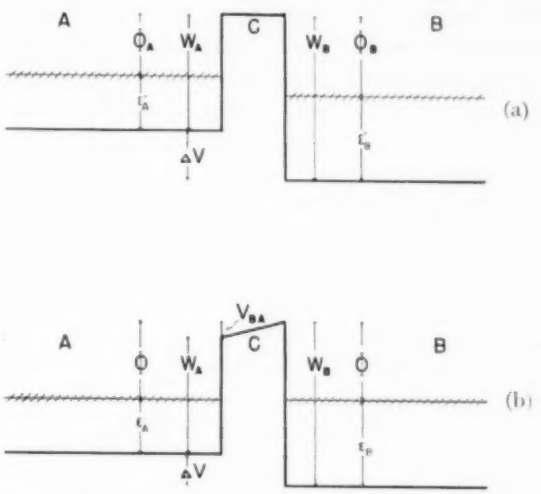


FIG. 4. (a) The potential energy of an electron in two uncharged pieces of different metals. (b) The potential energy of an electron in two pieces of different metals after equilibrium has been attained.

contact. The figures show the energies as a function of a single coordinate that is taken perpendicular to the surface of contact. In region A there is the metal A, and in region B there is the metal B. Region C represents the space between them, which may be treated as a vacuum. The heavy line represents the potential energy of an electron in these three regions, respectively. Figure 4(a) shows the situation between two uncharged pieces of metal. An electron in B has a lower potential energy than it has in A, and between them it has a constant potential energy, which may be taken as zero. The drop in potential energy that accompanies the passage of an electron through the surface is designated by W_A for metal A and W_B for the other. The hatched horizontal lines represent the energy of an electron that has a kinetic energy equal to ϵ_A' or ϵ_B' in the two metals. The hatching on these lines indicates that this is not a sharp upper limit for the occupied states; there are some electrons with kinetic energies above ϵ', and there are some states whose energies are less than ϵ' that are unoccupied. Nevertheless ϵ' is a perfectly definite kinetic energy that lies somewhere in the middle of this narrow limiting

region. It is, in fact, the kinetic energy part of the thermodynamic potential of the electron gas in the metal in question.

Figure 4(b) represents the situation when the two metals have been brought into contact and separated again, when they are connected at some other point, or when equilibrium has been set up by some other means. When the two metals really form parts of the same statistical mechanical system there can be only one value of the thermodynamic potential for the whole system. This does not mean that there can be only one ϵ' measured as a kinetic energy, but that there is an energy, kinetic plus potential, that is the same for all parts of the system. Let this be Φ, and let it be measured from a common zero of energy. The two hatched lines are now at the same level in the two metals. This is brought about because a few electrons from metal A have gone over to metal B and have raised the potential energy in B by enough to bring the tops of the two electron distributions together. The number of electrons necessary to do this is so small that it has no influence whatever on the electron densities in the two metals.

Between the metals there is now an electric field that corresponds to the potential difference V_{BA}. In the figure this is drawn as a difference in potential energy for electrons. The difference in the conventional electrostatic potential will have the opposite sign. This potential difference is the contact potential difference. If the two metals are touched together and then separated, B will be negatively and A positively charged, corresponding to this contact potential difference. The difference between the potential energies of electrons in the two metals will no longer be the same as when they were uncharged, but its amount can be determined from the relationships illustrated in the figure.

Further consideration of this figure shows that the contact potential differences cannot produce a current. It is true that the potential energy of an electron is less in one metal than in the other, but its kinetic energy is correspondingly greater, and the equality of the energies Φ in the two metals is just the condition for statistical equilibrium between them. In equilibrium there will be no tendency for electrons to flow one way or the other. On the other hand, the kinetic energy part of Φ is a function of temperature, and the temperature variation will be different for the two substances. This will produce a temperature variation in V_{BA} which might be supposed to be connected with the thermoelectric effects. But the constancy of Φ across the junction is the condition for equilibrium at any temperature, and there will be no tendency for a current to flow one way or the other across the junction. For this reason there will be no electromotive force that can be associated with the temperature dependence of the contact potentials.

Effect of an Electric Field

The discussion thus far has referred to the distribution of electrons over the various states in the absence of an electric field inside the metal and in the absence of a temperature gradient. The problem of the theory is, however, to find the distribution of the electrons under other circumstances, when both an electric field and a temperature gradient are present.

In the presence of an electric field along the x axis that tends to accelerate the electrons in the positive direction, the k_x of each electron increases uniformly and the whole distribution moves along the k_x axis in the positive direction. This uniform motion along the k_x axis is the representation of a constant acceleration in the x direction. Such an acceleration, and such a motion along the k_x axis, does not continue indefinitely. It is opposed by restoring forces

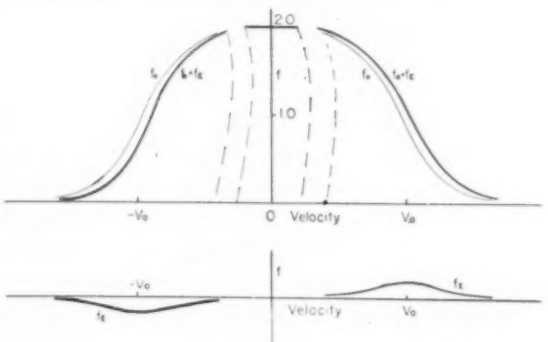


FIG. 5. Influence of an electric field on the electron distribution.

due to the collisions of the electrons with the ions of the lattice. Eventually there will be an equilibrium set up in which the distribution function of the electrons will have very much its original

shape, but will have its center displaced in the direction of positive k_x. The situation is indicated in Fig. 5 where the light line shows the equilibrium distribution in the absence of a field, and the heavy line shows the distribution when the field is present.

When the electrons are distributed as shown in Fig. 5 there exists a current. The number of electrons moving in the positive direction has been increased, and the number moving in the opposite direction has been decreased. One may also consider that an average drift velocity has been added to the random motion of each electron. This displacement under the influence of an electric field is very small. If it is such as to correspond to a velocity of 1 cm per sec., it represents a tremendous current, some 20,000 amp. per sq. cm, and such a shift is almost infinitesimal even on the scale of Fig. 5. It represents an increase in energy of the fastest electrons of about 1.2×10^{-19} erg or 7.5×10^{-7} electron volt. This is very much less than the temperature energy kT at the lowest attainable temperature, and to show it at all in Fig. 5, the displacement has been very much exaggerated.

The smallness of the change in the distribution due to the electric field justifies its treatment as a small correction to the equilibrium distribution. In Fig. 5 the equilibrium distribution is designated by f_0 and the correction due to the electric field by $f_E = Ee\Omega$. Because the shift along the k axis is small compared with the distance in which the distribution falls to zero, not only the shift but also this correction is proportional to E. The sum of these two quantities, $f=f_0+Ee\Omega$, corresponds to a current, and the magnitude of the current is given by the integral

$$I=\int ev_x f d\tau = E\int e^2 v_x \Omega d\tau = ES. \qquad (2)$$

$d\tau$ is the volume element in the space of k_x, k_y, k_z, and S is the electrical conductivity. S is a function of the temperature and of the metal. The integral involving f_0 vanishes because there is no current in the absence of a field. An approximate calculation of the current can be made by attributing to all of the electrons in f_E the speed v_0 corresponding to the kinetic energy ϵ'. For still higher order effects, however, such as the thermoelectric effects, it is sometimes necessary to take into account the fact that they do not all have quite the same velocity.

Effect of a Temperature Gradient

To investigate the thermoelectric phenomena it is necessary to study the effect of a temperature

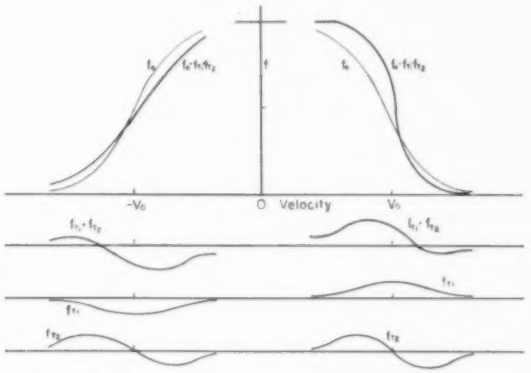


FIG. 6. Influence of a temperature gradient on the electron distribution. The temperature is increasing toward the right.

gradient on the electron distribution. Here again the effect is small and can be treated as a small correction to the equilibrium distribution. Figure 6 indicates the nature of the change in this case. The right side of the distribution is steeper, and the left side is less steep than in the absence of the temperature gradient. This can be associated qualitatively with the temperature gradient by considering that the electrons represented on the right are moving toward the right and have therefore come from the left. At positions to the left of the one in question the temperature is lower and the electrons coming from the region of lower temperature will reflect that fact. The electrons represented by the left part of the diagram have come from a position of higher temperature, and reflect that fact in their distribution. For the detailed derivation of the shape, reference must be made to some treatise on the theory of electrons in metals.[1]

In this case it is convenient to divide the correction to the equilibrium into two parts as shown in Fig. 6. The first part f_{T1} has the same shape as f_E, but in it Ω is multiplied by the temperature gradient instead of by the electric field. It is also multiplied by the rate of change

of ϵ' with respect to the temperature, and therefore f_{T1} is proportional to $d\epsilon'/dx$. Combining these terms, $f_{T1}=-(d\epsilon'/dx)\Omega$. If we consider at first only f_{T1}, it is clear that it represents a stream of electrons moving toward the high temperature end of the conductor. It is, of course, impossible to maintain such a state unless some means are available to introduce electrons at one end and remove them at the other. In an ordinary finite wire, the charge will accumulate at one end until there is sufficient electric field to correct the electron distribution in the opposite direction and to stop the current. Considering only f_{T1} the current will be

$$I=e\int v_x(f_E+f_{T1})d\tau=\int\left(Ee-\frac{d\epsilon'}{dx}\right)ev_x\Omega d\tau$$
$$=\left\{E-\frac{1}{e}\frac{d\epsilon'}{dx}\right\}S. \tag{3}$$

The current will be stopped when $E=(1/e)d\epsilon'/dx$. Associated with this opposing field there will be a potential difference between the two ends of the conductor equal to

$$\Delta V=-\int_1^2 Edx=\frac{1}{e}\{\epsilon'(T_1)-\epsilon'(T_2)\}. \tag{4}$$

The energy corresponding to ϵ' will be the same at both ends of the conductor, but the relative contributions of kinetic and potential will be different. Figure 7 shows the situation in which the temperature difference between the two ends is correlated with a potential difference.

Although this first part of the correction to the electron distribution will set up a potential difference between the two ends of a wire, and this potential difference will depend upon the material because $d\epsilon'/dT$ is different for different materials, such a mechanism will not produce any of the ordinary thermoelectric effects. To see that this is the case consider Fig. 8. Here it can be seen that the potential differences set up in the material due to the temperature gradient are just balanced by the potential differences that are set up at the junctions. The electric field at any point is proportional to $d\epsilon'/dx$ and this quantity, when integrated around any closed circuit will give zero. To the approximation considered here, the quantity Φ is constant throughout the whole circuit and this constancy guarantees that there will be no electromotive force, because it represents equilibrium of the electrons.

To understand the thermoelectric effects it is necessary to consider the next approximation and to include the term f_{T2} in the distribution function. This term is proportional to the temperature gradient, dT/dx, but it is independent of $d\epsilon'/dx$. It therefore exists in both metals where there is a temperature gradient, but it does not occur at the junctions, where ϵ' changes in a region of constant temperature. f_{T2} also contains $(\epsilon-\epsilon')/T$ as a factor, and this makes it a small quantity of higher order than either f_E or f_{T1}. The complete expression for f_{T2} is

$$f_{T2}=-\frac{\epsilon-\epsilon'}{T}\frac{dT}{dx}\Omega, \tag{5}$$

and the electric current associated with it is

$$e\int v_x f_{T2}d\tau=-\frac{1}{T}\frac{dT}{dx}\int ev_x(\epsilon-\epsilon')\Omega d\tau$$
$$=-\frac{1}{T}\frac{dT}{dx}N, \tag{6}$$

where the integral N is a function of the temperature and of the nature of the substance involved. Both the direction and magnitude of this current

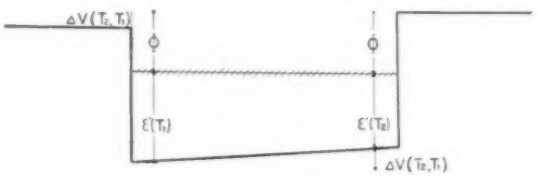


FIG. 7. Potential difference set up between the two ends of a conductor due to a temperature gradient acting through f_{T1} only.

are very sensitive to the exact value of the velocity associated with each electron state and also to the exact number of states per unit energy range.

For free electrons, f_{T2} represents a flow of electrons in a direction opposite to that of the thermal gradient, and therefore an electric current toward the high temperature end of the conductor. If all of the electron states repre-

sented a velocity of magnitude v_0, and the number of states per unit energy range were constant, f_{T2} would represent a zero current. However, both of these quantities vary considerably from metal to metal, and thus one can account for the wide variety of values for the thermoelectric coefficients.

The Thermal Electromotive Force

Because of the presence of f_{T2} and the existence of a current associated with it, the potential difference between the ends of a conductor in which there is a temperature gradient is not that given in Eq. (4). The potential difference will be either greater or less than this amount, depending upon the sign and magnitude of the current associated with f_{T2}. This means that the quantity Φ will not be a constant throughout the conductor in spite of its thermodynamic significance. This is possible because although Φ is a constant for a system in thermodynamic equilibrium, a metal in which there is a temperature gradient is not in equilibrium. It is carrying a heat current in the effort to set up such an equilibrium. The statistical distribution under these circumstances is not to be found by investigating the conditions for equilibrium, but must be found in other ways.

The extra potential difference due to the presence of f_{T2} is not compensated by any corresponding potential drop across the junction. When it is taken into account, the existence of a thermoelectromotive force can be understood as

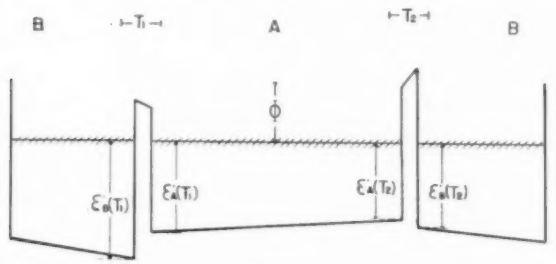


FIG. 8. Potential distribution set up in a thermoelectric circuit if only f_{T1} were effective.

resulting from the situation shown in Fig. 9. Its magnitude is given by the difference between the integrals of the fields in the two metals that serve to stop the current flow.

When all the correction terms are considered,

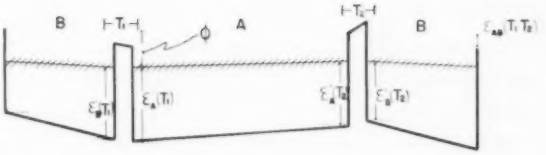


FIG. 9. Potential distribution in the thermoelectric circuit when the complete effect of the temperature gradient on the electron distribution is considered.

the total current is given by

$$I = e\int v_x(f_E + f_{T1} + f_{T2})\,d\tau$$

$$= \left\{E - \frac{1}{e}\frac{d\epsilon'}{dx}\right\}S - \frac{1}{T}\frac{dT}{dx}N. \qquad (7)$$

For the steady state on open circuit, when the current must vanish, it is necessary that the electric field be such as to cancel the other terms. This requires a field given by

$$E = -\frac{1}{e}\frac{d\epsilon'}{dx} + \frac{1}{T}\frac{dT}{dx}\frac{N}{S}, \qquad (8)$$

where both N and S are functions of the temperature and of the material. The thermal electromotive force is the integral of the negative of this field around the circuit. The first term gives no contribution to the integral because ϵ' is a single-valued function of position. The second term, however, does not vanish. It leads to

$$\mathcal{E}_{AB}(T_1, T_2) = \int_{T_1}^{T_2}\left\{\frac{N_B}{S_B} - \frac{N_A}{S_A}\right\}\frac{dT}{T}. \qquad (9)$$

This represents the integral of the "driving electromotive force" of Bridgman. It is located in the individual conductors and not in the junctions, and it is associated with the fact that the effort of the electrons to restore thermal equilibrium in the conductors leads them to flow along the conductor. This tendency is greater in one conductor than the other and so produces a resultant electromotive force. For a thermal electromotive force to exist in a circuit composed of pieces cut in different directions from a single crystal it is only necessary that N/S be different in different directions. S is known to depend upon direction in a crystal and it is to be expected that N will also depend on direction and in a different way.

Driving Electromotive Force

There has often been a good deal of confusion about the significance of the term "electromotive force." It is properly applied to the line integral, around some closed curve, of the vector force acting on a unit positive charge, and it is implied that such a force is proportional to the charge. If this total force, due to all possible sources, is designated by $\mathbf{E}_t$, then

$$\text{e.m.f.} = \mathcal{E} = \oint \mathbf{E}_t \cdot \mathbf{ds}. \tag{10}$$

It is clear that $\mathcal{E}$ depends upon the curve about which the line integral is taken, since the integral between two points is not generally independent of the path between them.

In addition to this use of electromotive force as a line integral around a closed curve, it is also possible to speak of the electromotive force along a specified path between two points, although this may be less customary. When this is done, the electromotive force between two points close together, along the shortest path between them, becomes equal to the scalar product of $\mathbf{E}_t$ and the distance between the points. This usage may lead to a confusion between $\mathcal{E}$ and $\mathbf{E}_t$, or between $\mathcal{E}$ and the potential difference between the points.

The total field $\mathbf{E}_t$ can be divided into two parts such that one part is the gradient of a potential. Let these two parts be $\mathbf{E}_s$ and $\mathbf{E}_d$, where $\mathbf{E}_s = -\text{grad}\, V$. $\mathbf{E}_s$ is then the field due to fixed electric charges, and because it is the gradient of a scalar potential, its integral around any closed curve will be zero. Then

$$\mathcal{E} = \oint \mathbf{E}_t \cdot \mathbf{ds} = \oint \mathbf{E}_s \cdot \mathbf{ds} + \oint \mathbf{E}_d \cdot \mathbf{ds}$$

$$= \oint \mathbf{E}_d \cdot \mathbf{ds}. \tag{11}$$

The electromotive force about any closed circuit is just the integral of $\mathbf{E}_d$ around the circuit, and the importance of the electromotive force is that it separates $\mathbf{E}_d$ from $\mathbf{E}_s$ in this way. The current in a circuit is due to $\mathbf{E}_d$ only, since $\mathbf{E}_s$ cannot maintain a continuous current. Bridgman[3] has designated this $\mathbf{E}_d$ as the source of the driving electromotive force of the circuit.

Working Electromotive Force

On the other hand, the electromotive force is often regarded as a source of energy, and when unit quantity of electricity moves around a circuit, the amount of work done on it will be $\mathcal{E}$. However, the electrons themselves have no more energy after completing the circuit than they had at its beginning. Along the path they have transferred their energy to the material of the conductor in various ways, and have transferred to it all of the energy gained. The electromotive force for the whole circuit, calculated from the energy relations, must be the same as the driving electromotive force for the whole circuit; but there is no reason for believing that they must be obtained by integrating the same fields. There is no reason for believing that the energy given to the electrons is always immediately transferred to the material. It may be carried along the circuit for a time, and then transferred to the material of the conductor at a distant point.

Bridgman defines a "working electromotive force" in terms of the energy given to a unit quantity of electricity when it moves around the circuit. Let $\mathbf{E}_w$ be a vector of the dimensions of force per unit charge and such that $\mathbf{E}_w \cdot \mathbf{ds}$ is the energy transferred from the material body to the electrons when one unit of charge moves a distance $\mathbf{ds}$. The assumption of such a vector implies that the rate of absorption of energy is proportional to the current and will be positive or negative depending upon the direction of the current. In addition to this energy exchange there is the Joule heat. The Joule heat represents energy transferred from the electrons to the material of the circuit to the amount of RI^2, where R is the total resistance of the circuit. If there are energy transformations proportional to other powers of the current they are neglected. The total rate of energy transfer to the electrons in the whole circuit is then

$$I \oint \mathbf{E}_w \cdot \mathbf{ds} - RI^2 = I \oint (\mathbf{E}_w - \mathbf{E}_s) \cdot \mathbf{ds} = 0. \tag{12}$$

This is set equal to zero because a group of electrons after completing the circuit have the same energy as at the beginning.

From Eq. (12) it follows that the integral of $\mathbf{E}_d$ around the circuit is the same as the integral of $\mathbf{E}_w$. The total electromotive force of one kind is the same as the total electromotive force of the other kind. It does *not* follow, however, that $\mathbf{E}_d$ and $\mathbf{E}_w$ are the same. In fact, in the thermoelectric case under consideration they are quite different, as was established by Bridgman from thermodynamic arguments. To evaluate $\mathbf{E}_w$ on the basis of the statistical theory, one must consider the flow of energy in the electron gas.

The heat current, or energy current, carried by the electrons is given by

$$W_0=\int v_x\epsilon f d\tau=\int v_x\epsilon(f_E+f_{T1}+f_{T2})d\tau$$

$$=\left(E-\frac{1}{e}\frac{d\epsilon'}{dx}\right)L-\frac{1}{T}\frac{dT}{dx}K, \qquad (13)$$

where

$$L=e\int v_x\epsilon\Omega d\tau, \quad K=\int v_x\epsilon(\epsilon-\epsilon')\Omega d\tau,$$

$$L=N+\epsilon' S/e.$$

The form of this expression shows that an energy current will exist in the presence of an electric field, a temperature gradient, a change of ϵ', or any combination of these three in which they do not neutralize each other. By means of Eq. (7) it is possible to eliminate the electric field, and to obtain an expression for the heat current in terms of the electric current and the temperature gradient. The result is

$$W_0=I\frac{L}{S}+\frac{1}{T}\frac{dT}{dx}\left\{\frac{NL}{S}-K\right\}. \qquad (14)$$

This gives the heat current entirely in terms of f_E when there is no temperature gradient, but in terms of f_{T1} and f_{T2} when no electric current exists.

The quantity that is of importance is not the flow of heat energy, but the accumulation of this energy at various places. Since the electron distribution is based upon the establishment of equilibrium due to interaction with the metallic ions, a change in the rate of heat flow represents an absorption of energy from the metal, or by the metal from the electrons. Let Q be the heat energy that accumulates per unit time per unit length of the circuit. This reduces to

$$Q=EI-\partial W_0/dx$$

$$=\frac{I^2}{S}-IT\frac{d}{dx}\left(\frac{N}{TS}\right)+\frac{d}{dx}\left(\kappa\frac{dT}{dx}\right), \qquad (15)$$

where κ is the thermal conductivity. The quantity Q is made up of two parts. The first part is the energy put into the electron gas by the electrostatic field E, and the second is the energy that is left at a given point because of the decrease in the energy stream. Strictly, this represents the accumulation of energy in the electron gas, but because the gas is in such intimate contact with the metal ions, and because it is always in thermal equilibrium with the metal ions, Q is the total deposition of energy at a point in the conductor.

The energy Q is expressed as a sum of three terms in the last form of Eq. (15). The first term is the ordinary Joule heat, and the last term is the accumulation of heat that would be attributed to a changing thermal conductivity. Neither of these is included in the idea of the working electromotive force. The working electromotive force is based on the integration of the coefficient of I around the circuit.

$$E_w=-T\frac{d}{dx}\left(\frac{N}{TS}\right). \qquad (16)$$

E_w is different from zero in a homogeneous metal in which there exists a temperature gradient, because of the presence of $1/T$ in the derivative, and also because both N and S are functions of the temperature. It is also different from zero at the junction between two metals because N and S depend upon the metal. The energy absorbed by the electron stream under the first condition gives rise to the Thomson effect. From its definition the Thomson coefficient is given by

$$\sigma_A(T)\frac{dT}{dx}=\frac{dT}{dx}\left\{\frac{1}{T}\frac{N}{S}-\frac{d}{dT}\left(\frac{N}{S}\right)\right\}. \qquad (17)$$

Under the second set of conditions the heat absorbed gives rise to the Peltier effect. From the

definition of the Peltier coefficient it follows that

$$P_{AB}(T)=\int_A^B E_w dx=-\int_A^B \frac{d}{dx}\left(\frac{N}{S}\right)dx$$

$$=\frac{N_A}{S_A}-\frac{N_B}{S_B}. \quad (18)$$

A given electric current flowing in metal A carries with it a different thermal current than the same electric current in metal B. The difference must be supplied at the junction and constitutes the Peltier heat. On the other hand, the thermal current accompanying a given electric current in a single metal depends upon the temperature, so that heat must be added or subtracted as the current flows toward regions of higher temperature. The exact values of these differences are very sensitive to the detailed properties of the electron states in the metals, and to the exact form of f_{T2}.

The Kelvin Relations

From the forms of Eqs. (17) and (18) it follows directly that

$$\frac{d}{dT}\left(\frac{P_{AB}}{T}\right)=\frac{\sigma_A-\sigma_B}{T}. \quad (19)$$

This is one of the Kelvin relations that was originally based on a thermodynamic treatment of the phenomena. It also follows from Eqs. (9), (17) and (18) that

$$\frac{d}{dT}\mathcal{E}_{AB}(T_1, T)=\frac{1}{T}\left\{\frac{N_B}{S_B}-\frac{N_A}{S_A}\right\}$$

$$=\frac{d}{dT}\left(\frac{N_A}{S_A}-\frac{N_B}{S_B}\right)$$

$$-T\frac{d}{dT}\left(\frac{N_A}{TS_A}\right)+T\frac{d}{dT}\left(\frac{N_B}{TS_B}\right)$$

$$=\frac{dP_{AB}}{dT}-(\sigma_A-\sigma_B), \quad (20)$$

which is the other Kelvin relation. It must not be concluded from the fact that these relations come out of the statistical treatment, that the statistical treatment provides a justification for the assumption of reversibility used in the thermodynamics. It is merely shown that *to the approximation of the statistical treatment* the Kelvin relations are valid. The whole statistical method is an approximation method. The electron distribution in the presence of the field and the temperature gradient is not obtained exactly, but is made up of the distribution without the field plus some small corrections. It is gratifying, although not at all surprising, that this treatment gives results similar to those obtained by Kelvin and Bridgman.

In his discussion of the thermoelectric phenomena, Bridgman worked out expressions for the quantities E_w and E_d by thermodynamic reasoning. According to him,

$$E_w=\sigma dT/dx \quad (21)$$

in the body of a homogeneous substance and the integral of E_w through a junction is the Peltier heat. This is identical with the result given in Eqs. (16) and (17) above. This identity is really only a question of the definition of E_w. Bridgman also works out the driving electromotive force as the integral of E_d, and he gives for E_d the expression

$$E_d=-\frac{dT}{dx}\left\{\int_0^T \frac{\sigma}{T}dT+f(T)\right\}.$$

This can be identified with the expression in Eq. (8) if σ/T and (N/TS) are taken as approaching zero as T approaches zero. This is to be expected both from Nernst's third law of thermodynamics and from the significance of the integrals N and S. It is also necessary that $f(T)(dT/dx)$ be identified with $-(1/e)d\epsilon'/dx$. This is not strictly a function of the temperature only; but it is a function of position in the circuit only, and gives a zero integral around the circuit. However, it has a value at the junctions as well as in the region of a temperature gradient in the homogeneous metal.

It can be seen from this discussion of the electron theory of the thermoelectric effects that the quantitative computation of the values of the thermoelectric coefficients is not a simple matter. It requires a very high precision calculation of the energy states for the electrons in a metal,

and it requires more precision in this calculation than is usually possible at the present time. Nevertheless it is possible to understand the qualitative features of the phenomena, the variety of magnitudes and signs and the temperature dependence, in terms of a picture that is identical with that used for the other properties of the metals.

Bibliography

1. For a detailed account see A. Sommerfeld, Zeits. f. Physik **47**, 1 (1928); A. Sommerfeld and H. A. Bethe, *Handbuch der Physik*, Vol. 24, Part 1, Edition 2; F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York and London, 1940) Chapter IV.

2. P. W. Bridgman, *The Thermodynamics of Electrical Phenomena in Metals* (Macmillan Company, New York, 1934).

3. Bridgman denotes E_d by $(\text{e.m.f.})_d$. It has been thought desirable here to distinguish between quantities whose dimensions are those of field strength, and quantities whose dimensions are those of field strength multiplied by length, by a more striking difference in notation.

Frictional Phenomena. I

This series of articles will be concerned with the present knowledge of viscous and related frictional processes, to be considered from a logical, unified standpoint. It will consist of three main parts covering, respectively, gases, liquids and solids. Theories, experimental data, and technical applications in the fields of mechanics, electricity, and chemical technology will be discussed.

BY ANDREW GEMANT

The Detroit Edison Company, Detroit, Michigan

Chapter I. Frictional Processes in General

IN this introductory chapter a general and rather formal theory of viscous processes is presented that is valid, with certain restrictions, for all three states of matter. In particular, the fundamental differential equation that can be used for the computation of any particular case is derived, and is then illustrated in the flow of a gas or liquid between two parallel walls.

1. Definition

By frictional phenomena are understood those surface processes that are involved in any relative motion of two bodies in contact with each other. This definition can be generalized by extending it to continuous media, any two neighboring elements of which being considered as the two bodies mentioned before. Frictional processes will then occur within any continuous medium if there exists in it a finite velocity gradient. These statements, and some more detailed ones to follow in the next section, are of a phenomenological nature. An explanation of the underlying physical processes for all three states of matter will be given in the three respective chapters.

Considering the fact that nearly any motion in nature can be described by relative velocities of neighboring elements as just mentioned, the great generality of the conception of friction will easily be understood. In fact, only astronomical motions, on the one hand, such as the rotation of the earth, for instance, and motions of atomic constituents such as the stationary rotation of electrons around nuclei, on the other, can be reckoned as performances without frictional accompaniment. But it is almost impossible to devise a motion in the intermediate range devoid of any frictional process.

If this be true for motions with which pure physics is concerned, the more will it hold for technical appliances of any kind. Thus, frictional processes will have a considerable bearing in the realm of the technical sciences, especially in applied mechanics, hydraulics, and electricity.

2. General Characteristics

The phenomena of friction show a great variety, as will be shown in detail in this monograph, yet it is desirable to point out here in the beginning the essential features that can be traced in any special type. In order to find this common feature, let us consider a rather typical case of motion involving friction. A solid body, the shape of which is not essential, should be assumed to float on water and to be dragged along with constant speed on the surface of a large water container.

The details of the velocity distribution of the water in the neighborhood of the floating body should not be considered. It is known that the water is carried along with the body at a speed that rapidly diminishes with increasing distance from the solid, and that, at a distance of several times the average dimension of the body, the water will be practically at rest. In any case, the water in the neighborhood will have a lesser speed than that of the body, and according to our definition given above, some frictional process will be operative between the two.

The chief result of this process will be a certain force exerted upon the body which is called the frictional force. The direction of this force is opposed to that of the speed of the body relative to the water. This is a general rule and will hold for any frictional process. Viewing viscous forces from this standpoint reveals them as forces that have to be overcome by the external forces in order to maintain a given motion.

According to the principle of relativity, the body can be treated as being at rest and the bulk of the water as moving in a direction opposite to that in which the body moved. Viewing the process from this angle, it follows that the body must exert a frictional force upon the water. As this force must be opposed to the relative velocity of the water, it must act in the direction of the motion of the body. The magnitude of the two forces is, of course, the same. This is just a special application of the general law of action and reaction. Looking at the viscous forces from this angle they are revealed as forces that result from a transmission of any motion, maintained by external forces, to other parts of a body. Both conceptions are instructive for the understanding of viscous processes.

Although the direction of frictional forces is determined as stated above, nothing of similar generality can be said with regard to the function expressing its magnitude. As a matter of fact, several of the different mathematical possibilities are encountered in nature. The most usual case is that the force is a function of the relative velocity itself. In this group again the more common type of relation is the simple proportionality between force and velocity. This was the case in our above example. However, any deviation from this linear law is realized in nature, as is the case with certain colloidal solutions, for instance.

The force might also be entirely independent of the characteristics of the motion, being just a constant. This is the case with some types of external solid friction.

Again, it is possible that the force, although entirely independent of the relative velocity, is a function of the displacement, reckoned from a certain position of equilibrium. Thus the internal frictional force, occurring during vibration in solids, is often a linear function of the amplitude of vibration. But, whatever the function might be, the sign of the force acting on a part is always opposite to that of the velocity of the part relative to its neighbor.

There is another general law which is a direct consequence of the first one. Let us return to our concrete example. In moving the body we have to perform a certain work which for the time unit will be

$$A_f = -uf,$$

if u = velocity, f = frictional force. This work is, according to our previous rule, always positive. Whatever the type of motion, A_f is positive in each of its elementary phases, and steadily accumulates. There is no motion imaginable, not even a purely vibrational one, moving the body to and fro through a position of equilibrium, in which even a part of the time integral $\int A_f dt$ should be zero. The reason for this is that f remains always opposed to u, whatever its magnitude.

The work A_f is totally converted into heat energy. This must be the case, since in a vibratory motion the whole system is exactly the same at the end as it was in the beginning. An accumulation of any other kind of energy is not possible, thus the work lost on the system can only turn into heat. In other words all frictional phenomena are of the dissipative type.

In mechanics, dissipative processes are treated in contradistinction to conservative ones in which latter there is an interconversion of work and kinetic or potential energy.

In thermodynamics frictional processes are called irreversible, as distinguished from reversible processes, such as the conversion of the internal energy of a system into external work and latent heat during an infinitely slow isothermal process.

3. *Deformation Velocities*

After this general review we proceed in formulating the equations that allow frictional processes to be treated mathematically.

As just pointed out, the expression representing the frictional force as a function of the characteristics of the motion is not universal.

This fact, unfortunately, prevents a generally valid theory from being formulated. There is, however, a way out of this difficulty, in considering that *the most frequent* relation is linearity between force and relative velocity. It is possible to formulate the equations under this assumption, and these equations will be presented in the following.

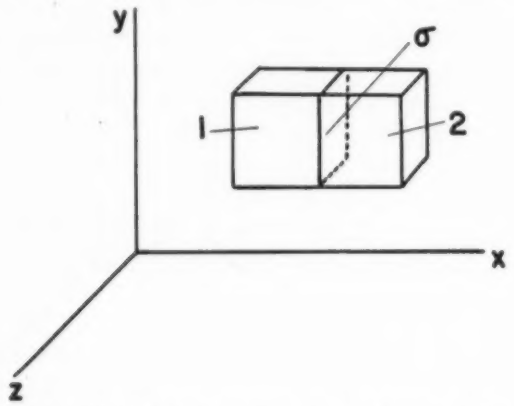


FIG. 1. Neighboring elements in continuous media.

One must, however, bear in mind that they are not generally valid, but mainly restricted to gases and liquids. For solids they cannot be applied without further comment. The reason for this is that molecular mechanisms underlying solid friction are entirely different from those encountered with gases and liquids.

Purely formally it is possible—and sometimes even useful—to apply these equations to solids as well, but those quantities that for gases and liquids represent characteristic physical constants will have only the significance of an auxiliary quantity, which is not a constant but might be any function of any other variables. We are referring here to the constant that will be defined later as the so-called viscosity constant of gases and liquids and which has a real physical sense. If introduced into cases dealing with solid friction, it is just a mathematical expression, and is some times referred to as the "equivalent viscosity."

We shall come back to this point more in detail in the third part of our monograph, dealing with solids, but it is important to mention it now in order that the scope of the following presentation may be clearly understood.

It is not our purpose to present a complete theory that is available in handbooks or textbooks; some essential points only should be given and explained, since they will be needed when dealing with special problems in the following chapters.

We consider a continuous medium (Fig. 1) and fix our attention on two neighboring volume elements (1, 2) possessing a velocity relative to each other, thus giving rise to frictional forces, as stated in Section 1. In other words, we imagine a certain surface element σ in the medium, separating the two volumes, and shall define the direction of the frictional force as being exerted from 1 to 2 along the surface σ. The surface might have any orientation with regard to the arbitrary coordinate axes x, y, z, but it is essentially sufficient to consider surfaces perpendicular to the three axes. If the behavior of the forces on these particular surfaces is known, then the construction of the forces on any other surface element is merely a problem of vector and tensor analysis.

Moreover, as we are mainly interested in isotropic bodies in which all directions are equivalent, it will be sufficient to consider one special surface element only, since the other two are perfectly analogous. Thus, we choose the surface perpendicular to the x axis and will put down expressions for the force acting from volume 1 to 2 along the surface unit of σ. Instead of the force, which is a vector, its three components should be considered. The x component is designated by X'_x. The prime should indicate that the force in question is a frictional one. The index x indicates that the surface in question is normal to the x axis, and that the force is transferred in the direction of the positive x axis; that is, from element 1 to element 2. In the same way Y'_x and Z'_x designate the y and z components executed upon a plane perpendicular to x.

All these forces should be linear functions of the relative velocities of neighboring particles. The velocity components of a particle with regard to the coordinate axes should be u, v, w. The relative velocities can then be expressed as the appropriate functions of the gradients of u, v, and w with regard to the coordinate axes.

If there is a flow in the y direction, for instance, then the velocity of element 2 in Fig. 1 relative to 1 is given by $(\partial v/\partial x)dx$, if the distance

between the centers of 1 and 2 is dx. Or, if u is taken into consideration, the relative velocity of 2 is $(\partial u/\partial x)dx$.

There are altogether nine such gradient components, describing the velocity distribution within the medium, but, as the detailed theory shows, there are only six quantities essential for calculating the frictional forces. The six quantities are:

$$x_x=\partial u/\partial x; \quad y_y=\partial v/\partial y; \quad z_z=\partial w/\partial z;$$
$$x_y=\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}; \quad y_z=\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z}; \quad z_x=\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}, \tag{1}$$

the expressions x_x, x_y, etc., being nothing else but shorthand writings for the respective gradient components. Thus, the nine components occur only in the six combinations given by Eq. (1). Owing to the symmetry of the expressions, it is just as well to write y_x for x_y, z_y for y_z and x_z for z_x.

The six expressions (1)—to be exact, the quantities: x_x, y_y, z_z, $\frac{1}{2}x_y$, $\frac{1}{2}y_z$, $\frac{1}{2}z_x$—are termed the components of the tensor of the *deformation velocity*. The name originates from the theory of elastic media, and it will be quite useful to illustrate the correctness of this designation. Let us consider the term $\partial u/\partial y$. A plane $ABCD$ of the body (Fig. 2) situated in the xy plane of the coordinate system should assume the shape $A'B'CD$ after a certain flow in the x direction has taken place. The displacement of any point should be denoted by a, being a function of y as well as of the time, t.

Now the velocity of each point is $u=\partial a/\partial t$, and the gradient of this velocity as considered above is $(\partial/\partial y)(\partial a/\partial t)$.

Since differentiation with regard to y and t is independent, their sequence can be altered, and the gradient written as $(\partial/\partial t)(\partial a/\partial y)$. The gradient in the bracket is a measure of the deformation which in this special case is a simple shear. The whole expression is then the change of shear with time, or the deformation velocity, also called rate of shear in this case. Hence, it is shown that the conceptions of velocity gradient and deformation velocity are equivalent. This result is of importance in considering elasto-viscous materials.

According to our fundamental assumptions, the force components $X'_x\ldots$ are linear functions of the deformation velocity components $x_x\ldots$ and the question is how many proportionality constants will be required. A consideration of

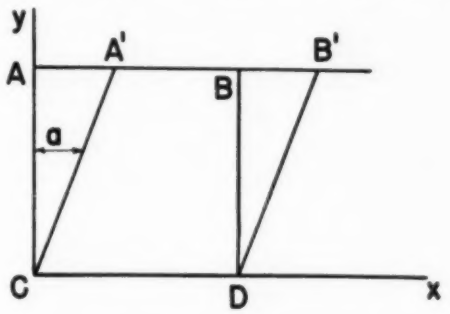


FIG. 2. Diagram of deformation.

energy relations in conjunction with requirements arising from the isotropy of the body shows that generally two constants are sufficient to describe the frictional behavior of an isotropic medium. Whether both are necessary is another question, and the generally adopted view is that there must be some simple relation between the two, leaving only one characteristic constant for an isotropic medium.

The linear equations to be obtained by means of the energy relations just mentioned are the following:

$$X'_x=-2\eta x_x-\xi(x_x+y_y+z_z)$$
$$Y'_x=-\eta y_x; \quad Z'_x=-\eta z_x, \tag{2}$$

the two constants being η and ξ. As mentioned before, similar equations are valid for the planes normal to the y and z axes. Although the deduction of (2) is omitted here, a kind of interpretation in two special cases will be given in the next section.

4. Stress Components

Before explaining Eq. (2), the relations between η and ξ should be put forward. The sum $(x_x+y_y+z_z)$ means, as is known from vector analysis, the total expansion of the volume unit per second, as is designated by div$\mathbf{U}$, if $\mathbf{U}$ is the velocity vector at any point.

Further, one has to remember that in liquids and gases there is always present the hydrostatic pressure p which acts in the same direction as X'_x, namely, normal to any surface. Thus, one can express the total surface forces as the sum of the frictional forces and the hydrostatic

pressure, and designate them as X_x, Y_x, and Z_x. One has then,

$$X_x = p - 2\eta x_x - \xi \operatorname{div}\mathbf{U}$$
$$Y_x = -\eta y_x; \quad Z_x = -\eta z_x. \tag{3}$$

For incompressible liquids evidently

$$\operatorname{div}\mathbf{U} = 0$$

and, from the first Eq. (3):

$$p = \tfrac{1}{3}(X_x + Y_y + Z_z),$$

the equation for Y_y and Z_z being similar to that for X_x. For incompressible liquids the hydrostatic pressure proves to be the mean value of the three normal force components.

This rather plausible consequence should now be assumed to hold for compressible liquids (i.e., gases) as well, an assumption that is in keeping with experimental observations. Carrying out the addition, one sees that the condition necessary for this assumption to hold is:

$$\xi = -\tfrac{2}{3}\eta.$$

Whereas, therefore, the first of Eqs. (3) for liquids assumes the final form

$$X_x = p - 2\eta x_x, \tag{4}$$

that, for gases is

$$X_x = p - 2\eta x_x + \tfrac{2}{3}\eta \operatorname{div}\mathbf{U}. \tag{5}$$

The only constant remaining in the equation is now η; it is characteristic of the medium in question and depends only on such variables as pressure and temperature. It is called the viscosity constant, and will be considered more in detail in several of the following chapters.

As to the interpretation of Eq. (3), let us first consider the case wherein of the nine velocity gradient components only one, $\partial u/\partial y$, is finite and positive; all the others are zero. Eq. (3) then reduces to

$$X_x = Y_y = Z_z = p$$
$$X_y = Y_x = -\eta \partial u/\partial y. \tag{6}$$

The other force components are zero. Figure 3(a) shows at the bottom the distribution of u with regard to the $x-y$ axes, and above, a volume element. The particular type of motion, as assumed, is maintained by certain external forces, acting at the boundaries of the body, and the frictional forces must be such that they transmit the particular motion, in this case a pure shear, into the interior of the body. This is done exactly by the forces parallel to the x axis, as given by Eq. (6) (X_y).

Simultaneously, however, there are also forces parallel to the y axis, as required by Eq. (6), and shown in the figure. The significance of this latter is the following.

X_y is in reality a couple, trying to rotate any volume element in a clockwise direction. Y_x, on the other hand, produces a couple in an anticlockwise direction. In every case of a stationary flow those two opposite couples will exactly compensate each other; the resulting torque is zero and, should the volume element rotate, which generally is the case, then its momentum of rotation must remain constant.

In the second case let us consider $\partial u/\partial x$ finite and positive, all other components vanishing. Then Eq. (3), with the value introduced for ξ, becomes

$$X_x = p - \frac{4}{3}\eta \frac{\partial u}{\partial x}; \quad Y_y = Z_z = p + \frac{2}{3}\eta \frac{\partial u}{\partial x}, \tag{7}$$

the other components being zero.

Figure 3(b) illustrates this case. The velocity gradient is shown at the bottom. Above is a volume element with the frictional forces ($X'_x \ldots$). Since the motion as assumed leads to an expansion of the medium, there must be a negative pressure component in the x direction as required by the first Eq. (7). In addition, there are positive pressure components in the y and z direction, the magnitude of which is such that the sum of these three friction components should be zero, or the sum of the total components should be $3p$.

The consideration of these two special cases should suffice as an interpretation of Eq. (3), since all the others are different combinations of these two elementary ones.

5. *The Fundamental Differential Equation*

In dealing with experimental cases involving friction another step is necessary. Up to now we have formulated the frictional forces as acting on surface elements. This is, indeed, correct and in keeping with the physical picture according to which the force is transmitted from

one body to another across their common boundary, or from one volume element to another in a continuous medium. If, however, the velocities of these volume elements have to be calculated—and this is one of the most frequent problems—then these frictional surface forces have to be transformed into volume forces. This is a frequent operation in mathematical physics, and we will consider how it is done in our special problem.

A cube of volume one, with edges parallel to the coordinate axes, should be considered within the medium in question. There are generally four types of forces acting upon it.

(a) Let us begin with the frictional ones. The six forces acting upon its six surfaces will evidently be equivalent to the forces acting on the volume. In summing them up, it will be useful—following the procedure adopted in the previous sections—to separate the vectors in question into their three components. Only the x components should be dealt with.

The x components of the frictional forces acting on the yz planes are X'_x acting on the one surface and $-X'_x$ acting on the other. Were X'_x constant throughout the medium, the two would just compensate each other. Only in case there is a finite gradient of this force, will the resultant differ from zero. One of the forces

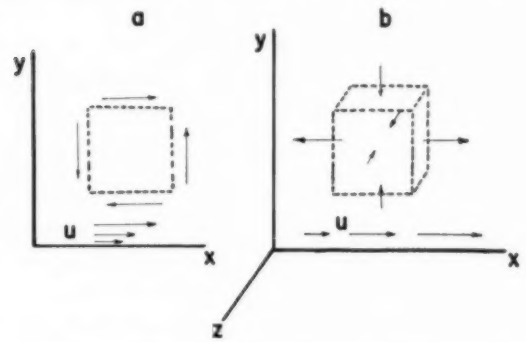


FIG. 3. Diagram of stresses: (a) $\partial u/\partial y$ finite; (b) $\partial u/\partial x$ finite.

is then X'_x, the other $-(X'_x+\partial X'_x/\partial x)$, since the cube has been taken as of volume one, and their sum is $-\partial X'_x/\partial x$, acting in the positive x direction.

The same situation holds for the X'_y and X'_z components. The actual force acting on the volume is always given by their respective differential quotients. Thus, the total frictional force is

$$-\frac{\partial X'_x}{\partial x}-\frac{\partial X'_y}{\partial y}-\frac{\partial X'_z}{\partial z}.$$

(b) In the same way the hydrostatic pressure has to be taken into consideration. Only in case of a pressure gradient is a force exerted upon that volume element; its magnitude is $-\partial p/\partial x$. The forces under (a) and (b) can be dealt with together, by replacing X' with X.

(c) There might be any external force acting on the volume elements, such as gravitation or electric forces in case of free electric charges, the x component of which should be designated as F_x.

(d) For nonstationary processes the velocity is not constant with time. The inertia force has then to be considered too, having an x component $-\rho\partial u/\partial t$, if ρ is the density.

Adding the forces reviewed under (a) to (d), the following equation of motion valid for each volume element is obtained:

$$F_x-\rho\frac{\partial u}{\partial t}-\left(\frac{\partial X_x}{\partial x}+\frac{\partial X_y}{\partial y}+\frac{\partial X_z}{\partial z}\right)=0 \qquad (8)$$

and exactly analogous ones for the y and z components.

6. *Energy Relations*

At the end of this introductory chapter the main energy relation controlling frictional processes should be given, such as can be derived on the basis of the above theory. The same limitations hold with regard to solid friction, as indicated in Section 3.

The derivation of the equation will not be given; we again content ourselves with having it explained.

In the last section external forces F have been introduced which, together with the hydrostatic pressure p, can be considered as the driving forces. They will perform a certain external work per second, the amount of which should be A. This work will be used up in three different ways.

First, it might be transferred into the *kinetic energy* L of the system. This transformation will take place during transitory, or initial

processes, and the increase per second of the kinetic energy should be denoted by ΔL. This quantity is evidently zero for stationary motions.

Second, it might be stored as the elastic energy U of the system, playing a role only in gases possessing a finite compressibility and in solids with finite elasticity (also elasto-viscous liquids—see the corresponding chapter). Its increase per second is ΔU.

Third, a certain amount of heat is developed due to friction which, as has been shown, is a dissipative process. The heat produced per second and per volume unit of the medium should be W.

Then we have the equation

$$A=\Delta L+\Delta U+\int W dv, \tag{9}$$

the last integral being taken over the total volume.

Our special problem is now to express W as a function of the force and deformation velocity components introduced into our general theory. This function is the following:

$$W=-(X'_x x_x+Y'_y y_y+Z'_z z_z+X'_y x_y + Y'_z y_z+Z'_x z_x). \tag{10}$$

Instead of a rigorous proof which is outside the scope of this monograph, an interpretation should be given similar to that of Eq. (3) in Section 4.

Take $\partial u/\partial x$ as finite, all other velocity gradient components vanishing, and consider again a cube of unit volume as in Section 5. Equation (10) then reduces to

$$W=-X'_x \partial u/\partial x, \tag{10a}$$

where X'_x acting on the two yz planes is given by Eq. (2). Since the motion of the cube in question is relative to the surroundings, the left of the two yz planes (see Fig. 3(b)) should be considered for the moment as fixed in space and only the right one moving. Its velocity relative to the space and in the positive x direction is then $\partial u/\partial x$. X'_x is the force acting upon the volume unit on this right surface, being directed toward the negative x direction. Its magnitude in the positive x direction is then $-X'_x$. The work done per second by this force is then, according to elementary mechanics, $-X'_x \partial u/\partial x$ (force times path). Since, however, this work must be converted entirely into heat (Section 2), we have exactly the above equation.

Again, let us take only $\partial u/\partial y$ as different from zero. Equation (10) becomes

$$W=-X'_y \partial u/\partial y. \tag{10b}$$

X'_y, acting on the $x-z$ planes, is given by (2). The lower $x-z$ plane (Fig. 3(a)) should now be fixed, the upper moving with a velocity $\partial u/\partial y$ in the positive x axis. The force acting on this plane and in this direction is $-X'_y$, its work per second $-X'_y \partial u/\partial y$, as Eq. (10b) requires.

In summing up over all gradients in the same way, one obtains the complete form given by Eq. (10).

It is, of course, possible to substitute Eq. (2) into (10), and obtain expressions, which sometimes are more advantageous. Thus, (10b), for instance, becomes:

$$W=\eta(\partial u/\partial y)^2, \tag{10c}$$

showing that, whatever the sign of the gradient, the developed heat is positive.

It is important to note that, although in some cases the actual external forces are restricted to certain boundaries of the medium in question, yet heat is being developed in each volume element of the medium possessing a finite velocity gradient. Certain cases will be dealt with in the next chapters, illustrating this behavior.

7. Flow Between Parallel Walls

A special case should be considered here, in order to give an application of the equations deduced. Two parallel walls in a distance of $2h$ (Fig. 4) should form a channel in which a flow of the velocity u in the direction x takes place. The driving pressure difference per unit length should be p. The height of the channel (y direction) is unessential. All data below will refer to a height of unity.

Generally we have to assume that u is a function of z. Thus, writing down the components of the deformation velocities (1), we have only

$$x_z=du/dz,$$

all others being zero.

As to the stress components (3), we have according to Eqs. (6):

$$X_z = -\eta \frac{du}{dz}; \quad X_x = Y_y = Z_z = p.$$

Now we apply the differential equations (8) to a volume element of thickness dz as shown in the Fig. 4. There are no volume forces F present, and in the steady state in which we are interested, the inertia forces disappear. Consequently, we have for the x component

$$\eta \frac{d^2u}{dz^2} - \frac{dp}{dx} = 0,$$

whereas the other two components yield

$$\frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} = 0,$$

indicating that a pressure difference acting in the x direction is the only one. Denoting $-dp/dx$ by p as mentioned above, and integrating, we have

$$du/dz = -pz/\eta, \tag{11}$$

the integration constant being zero, since, for reasons of symmetry, du/dz is zero for $z=0$. For the next integration the boundary condition at the wall has to be settled; this question will be dealt with in the next chapter. Let us here assume the most frequent case, namely, that there is no slip at the wall: $u=0$ for $z=h$. Then

$$u = (p/2\eta)(h^2 - z^2). \tag{11a}$$

The velocity distribution along the channel thickness is a parabolic function with its maximum at the center plane.

The total volume V flowing through the unit height is given by

$$V = \int_{-h}^{+h} u dz = \frac{2ph^3}{3\eta}, \tag{11b}$$

being proportional to the third power of the channel thickness.

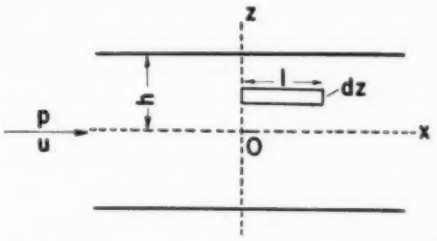


FIG. 4. Diagrams of flow between parallel walls.

The heat produced is given by Eq. (10). Using (10c) for the volume element dz, we have

$$dW = \eta x_z^2 dz = \frac{p^2 z^2}{\eta} dz \tag{11c}$$

and

$$W = \int_{-h}^{+h} dW = 2p^2h^3/3\eta, \tag{11d}$$

which, as can be seen, yields the relation

$$W = pV. \tag{12}$$

In other words, the heat developed is equal to pressure gradient times flow. The resistance of the channel, R, per unit length can be defined by means of

$$R = p/V, \tag{13}$$

thus, giving in this case

$$R = 3\eta/2h^3. \tag{11e}$$

The simple case discussed is most instructive in clarifying the theory of viscosity. Similar cases, which are more important from the standpoint of experimental devices, will be treated in some of the following chapters.

The Theory of the Plastic Properties of Solids.* IV

BY FREDERICK SEITZ

Randal Morgan Laboratory, University of Pennsylvania, Philadelphia, Pennsylvania

AND

T. A. READ

Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania

Part B. Polycrystals[1]

INTRODUCTORY DISCUSSION

THE properties of polycrystals are influenced by two separate factors—first by the intrinsic properties of the single crystal constituents or grains, and second by the restrictions neighboring grains exert on one another. If it were not for the second influence, all properties of polycrystals could be derived by taking some simple average of the properties of corresponding single crystals for various orientations. This simple procedure is possible in a few simple cases, of course. For example, the elastic moduli and electrical resistivity of pure polycrystals may be approximated closely[2] by taking the appropriate averages for single crystals. On the other hand, many properties of polycrystals, such as shear strength and internal friction, are affected by the discontinuities in structure that occur in a way that cannot be explained simply by treating the system as a set of disoriented isolated single crystals. We shall refer to the additional factor as *grain boundary influence* and shall attempt to unify present knowledge of this factor in the next section. It should be borne in mind that by grain boundary influence we mean not only the effect arising from the atoms in the transition region between two grains, but also the effect that neighboring grains exert on one another.

On the whole, the amount of purely scientific curiosity that has gone into the experimental study of polycrystals is far less than that expended in the investigation of single crystals. This fact will be evident at many points in the following pages, for only in a few cases is it possible to draw conclusions from available experimental work comparable with those that may be drawn from the basic experiments on single crystals.

7. GRAIN BOUNDARY INFLUENCE

a. The grain boundaries as barriers for thermal and electrical flow

Since different grains presumably grow from different nuclei, it follows that grain boundaries will naturally be the regions where insoluble impurities will tend to aggregate. For this reason, the grain boundaries may act as barriers for thermal and electrical conductivity, particularly in more impure materials. We shall see later that this fact has an influence on the internal friction of polycrystals.

b. The nature of the transition layer between grains

Even if the material of which polycrystals are made is very pure, or if all impurities are highly soluble, we may expect an abnormal arrangement of atoms in the immediate vicinity of the boundary between grains. That is, we may expect a transition layer of atoms which occupy positions resulting from a compromise between the forces of the atoms in both grains. These atoms will not be so tightly bound as the atoms in the interior of the grains and, as a result, may be expected to be more mobile at a given temperature than interior atoms. This does not necessarily mean that slip will occur more easily along grain boundaries than within grains for, as we have seen in the previous sections, slip is ordinarily determined by the ease of formation

* The previous installments of this series appeared in the February, March, and June, 1941, issues of this journal.

[1] The analytical conditions used in mathematical treatments of practical problems in plasticity will not be surveyed here. They are discussed in the book by A. Nadai, *Plasticity* (McGraw-Hill Book Company, New York, 1931), and that by M. Gensamer, *Strength of Metals Under Combined Stresses* (American Society for Metals, Cleveland, 1941).

[2] See, for example, E. Schmid and W. Boas, *Kristallplastizität* (Springer, Berlin, 1936), Section 81.

and motion of dislocations rather than by the ease with which individual atoms may move. At most we may conclude that the diffusion of a given atom would take place more easily[3] at the grain boundaries than inside of grains. In fact, it is possible that, as a result of irregularities occurring in the intergranular regions, it would be difficult for dislocations to move through them.

The width of the transition region between grains in very pure crystals is not known from experimental work. We know from other sources, however, that the forces between atoms in solids are of short range, extending with appreciable intensity only over a few atom distances. For this reason, it seems safe to conclude that the width of the intergranular region is of the order of five interatomic distances at most. This conclusion is not valid, of course, in materials containing a high percentage of insoluble impurities, for in such cases a large fraction of this material may be localized at the grain boundaries.

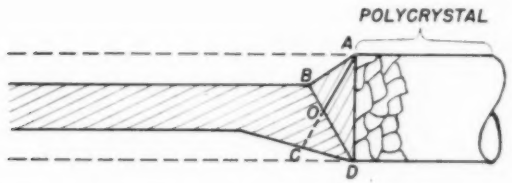


FIG. 51. Slip in a single crystal of zinc bounded by a polycrystal. (After Miller.) The heavy contour is the shape of the specimen after slip, the dotted line the shape before. The line *AD* marks the boundary between the single crystal and polycrystal. Slip has not occurred in the region *OAD* neighboring this boundary. The light lines in this region designate the position of the latent slip planes, and show the orientation of the basal plane in the entire single crystal before slip. Simple slip has occurred in the regions *OBA* and in the part of the specimen lying to the left of *BD*, but the latter region has been bent relative to the plane *BD*. The light lines show the slip bands in this region. It may be noted that slip occurs by bending in the region *OCD*, which contains planes previously intersecting *AD*.

From a study of the behavior of very pure tin at temperatures near its melting point, Chalmers[4] has concluded that the grain boundary material has a slightly lower melting point than the bulk material. This conclusion was drawn from the fact that grains separate along their boundaries at temperatures somewhat below the melting point. The difference between the separation temperature and the true melting temperature for any pair of grains turns out to be independent of the relative orientation of the two crystals and of the amount of impurity, providing it does not exceed 0.02 percent. For high purity tin the measured temperature difference was 0.14°C. These results are apparently in good accord with the conclusion drawn previously that the transition layer of atoms is thermodynamically less stable than interior atoms.

There is considerable evidence that the bond at grain boundaries is very strong at temperatures not too near the melting point, in spite of this lower thermodynamic stability. For example, fracture occurs most commonly[5] *through* grains rather than at their boundaries in rupture tests well below the melting point. On the other hand, intercrystalline fracture is common near the melting point. At first sight, observations of this kind seem to support the view that the intergranular material actually is stronger than the bulk material at low temperatures and weaker at high temperatures. However, an alternative explanation of these facts is as follows. At low temperatures the grain boundary material is somewhat weaker than the grains and rupture starts at grain boundaries. The difference in strength is not so great, however, that a crack will automatically follow grain boundary surfaces regardless of their inclination relative to the plane of greatest tensile stress. In fact, once started, a crack will occur in the plane of greatest tensile stress even if this plane cuts through grains. At high temperatures it is possible that the greater relative instability of grain boundaries increases and a crack will follow such boundaries. Indeed, Chalmer's experiments indicate that the grain boundaries are extremely weak just below the melting point of the bulk material.

The observation that low temperature fracture is principally transcrystalline rather than inter-

[3] It does not necessarily follow that bulk diffusion occurs more rapidly along grain boundaries, for the total amount of intergranular material is probably very small in a pure polycrystal.

[4] B. Chalmers, Proc. Roy. Soc. **A175**, 100 (1940).

[5] See, for example, the discussion in the book by Z. Jefferies and R. S. Archer, *The Science of Metals* (McGraw-Hill Book Co., Inc., New York, 1924), p. 67.

granular was first explained by Beilby[6] on the assumption that each grain in a polycrystal is surrounded by a thick coating of brittle amorphous material having much greater strength than the interior of the grains at temperatures not too near the melting point. Apparently this notion must be almost completely discarded, for investigation[7] of the electrical behavior of evaporated films of many common metals indicates that crystallization can occur at temperatures near 100°K. Thus amorphous coatings many atom layers thick would be highly unstable under ordinary conditions.

More direct evidence against a thick transition layer hypothesis has been given by Chalmers[6] from a study of the critical shearing stresses in bicrystals of tin. He investigated cases in which the axis of tension was symmetrically disposed relative to the two grains in the specimen, so that the stresses were the same in corresponding crystallographic planes in each grain. Results showed that the critical shearing stress increased continuously as the relative orientations were changed, starting from the case in which they were the same (i.e., a single crystal). Thus he concluded that there is no essential discontinuity in the nature of the grain boundary in passing from a single crystal to a polycrystal, as would follow if the thick transition layer theories were correct. The origin of the increase in shearing stress with increasing disorientation will be discussed below.

c. Grain boundaries as the seat of stress magnification and dislocations

If block boundaries in single crystals can contain weak spots for stress magnification, as experiments on slip and rupture indicate, we may expect grain boundaries to have similar properties. Thus we may expect grain boundaries to be the source of dislocations. Excellent evidence for this may be derived from experiments on internal friction of the type discussed in Section 2 Part *d*. If single crystals and course-grained polycrystals of very pure copper are carefully annealed, both have very low decrements, of the order of 10^{-5}.

Now we saw in Section 2 that, in the case of single crystals, this type of internal friction is very sensitive to mechanical treatment, presumably because the number of dissipating centers (dislocations) is increased as a result. It is found that the coarse-grained polycrystals are much more sensitive. For example, the internal friction of an ordinary sized single crystal of copper is scarcely affected when the specimen is dropped on a wooden surface from a height of one inch, whereas the internal friction of an otherwise identical polycrystal is raised by a factor of about ten. This result indicates that grain boundaries are a ready source of dislocations and implies, in turn, that slip nuclei are easily formed there. Whether this is due to the fact that there are larger numbers of weak spots at grain boundaries or to other causes cannot be said at present.

d. Slip interference at grain boundaries

Neighboring grains exert a strong restricting influence on the amount and kind of slip that

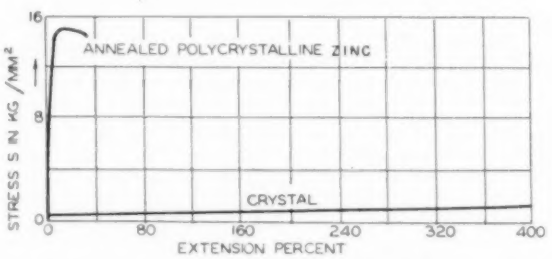


FIG. 52. Comparison of the stress-strain curves for single crystal and polycrystalline specimens of zinc. (After Elam.)

may occur within the grains. A significant experimental investigation of this effect has been carried out by Miller[8] on specimens of zinc consisting of a large single crystal bounded by a polycrystal (Fig. 51). Measurements were made on specimens extended at 180°C since slip occurs smoothly at this temperature. Miller found that those slip planes of the single crystal that intersect the polycrystalline region cannot operate as freely as those which do not intersect. Thus in the typical case illustrated by the figure,

[6] A survey of the theories of intercrystalline bonding is given by E. H. Bucknall, Metals Industry **311**, 369, 396 (1929). See also B. Chalmers, Proc. Roy. Soc. **A162**, 120 (1937).

[7] R. Suhrmann and G. Barth, Physik. Zeits. **36**, 841 (1935); Zeits. f. Physik **103**, 133 (1936); R. Suhrmann and W. Berndt, Zeits. f. Physik **115**, 17 (1940).

[8] R. F. Miller, Trans. A. I. M. E. **111**, 135 (1934).

stresses sufficient to cause slip in the uninhibited regions of the crystal were not sufficient to cause slip in the region *OAD* immediately bordering the polycrystal. Although slip occurred in the region *OBA* which does not border the polycrystal, the deformation was highly restricted by the presence of the unstrained region in that rotation of the slip planes was not permitted. Similarly, slip occurred in the region *OCD* but was limited to the amount allowed by bending on the plane *BD*. Miller states that the total slip in this region is somewhat less than that occurring in the unrestricted part of the single crystal.

The fact that slip can occur in region *OCD* makes it seem surprising, at first sight, that slip is absent in *OAD*, for if the shearing stress there had the same value as in the rest of the specimen, we might expect some bending, as in *OCD*. A reasonable explanation of this fact is as follows. Initially, the shearing stress is uniform throughout all slip planes (exclusive of microscopic fluctuations), but after an undetectably small strain has occurred in *OAD*, the stress becomes redistributed in such a way that the shearing stress is lower in this region. This evidently requires that the polycrystalline region exert a transverse tensile stress on the material to the left of the boundary *AD*, thereby reducing the shear in the slip planes. A condition sufficient for this is that slip does not occur in the intergranular surfaces along *AD*. Naturally, the transverse tension across *AD* will diminish if the polycrystalline material becomes deformed, but it is clear in any case that the slip occurring in *OAD* and in the polycrystalline region must be closely correlated.

Now if a single crystal grain is favorably oriented for slip and is entirely surrounded by less favorably oriented grains, as occurs in the interior of a polycrystalline specimen, stresses similar to those occurring in Miller's cases will be exerted across the boundaries of this grain and slip will be permissible only when the stress is sufficient to allow all grains to deform. Naturally, the more numerous and more randomly oriented the planes of easy slip are, the more probable it will be that one of the easy slip planes of an arbitrarily chosen grain is favorably oriented for deformation. Thus the stresses

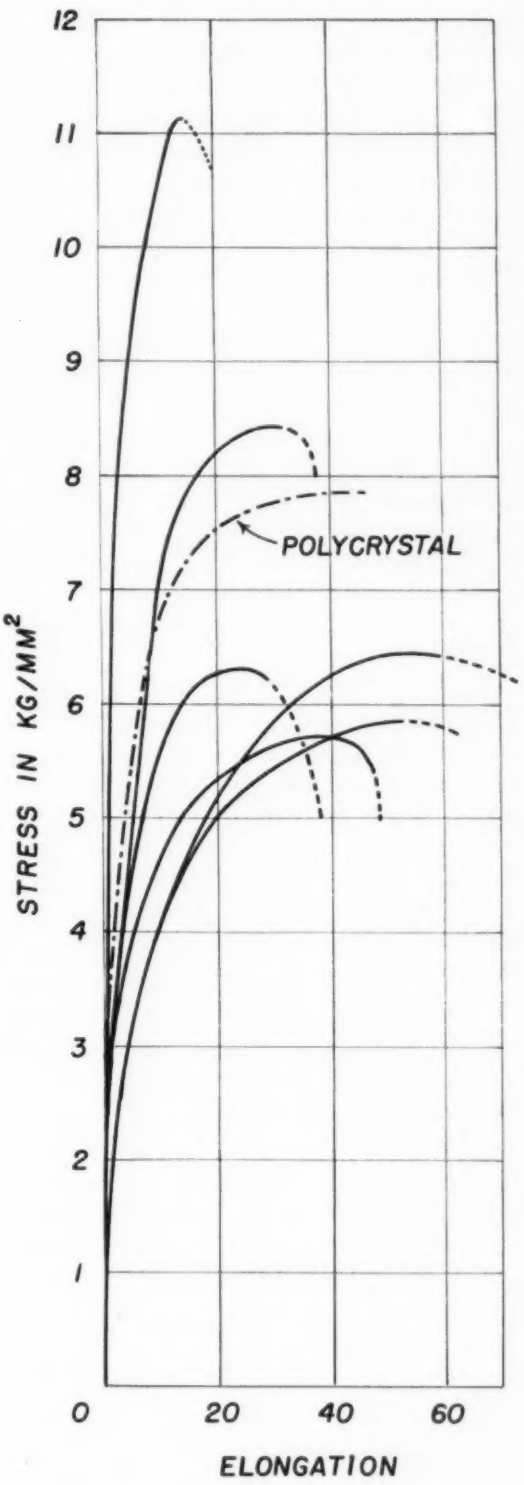


FIG. 53. Same as Fig. 52 for aluminum. In this case the polycrystalline curve is intermediate between extremes of the curves for variously oriented single crystals. (After Schmid and Boas.)

needed to produce a given deformation should be smaller for cubic crystals, which have many easy slip planes, than for hexagonal crystals, such as zinc, which have only one. Figures 52 and 53 bear out this conclusion by showing that the difference between the stress-strain curves for single and polycrystalline specimens is much greater in zinc[9] than in aluminum,[10] which is a face-centered cubic crystal possessing four equivalent planes of easy slip.

It is natural to ask at this point whether or not the actual stress-strain curves for polycrystals can be explained quantitatively by regarding them as a cluster of single crystals subject to a simple constraint, such as that there shall be no relative motion at grain boundaries. This problem has not yet been subject to careful analysis so that it is not possible at present to say that the constraining condition is a simple one. We shall return to this topic briefly in the next section.

One of the facts indicating that the strength of polycrystalline specimens cannot be derived with the use of a simple constraining condition is that the Brinell hardness,[11] the form of the stress-strain curves, and the breaking strengths of polycrystals seem to vary with grain size[12] in a range in which the grain size is much smaller than the dimensions of the specimen and of the ball used in the test. For example, Fig. 54 shows the variation of the hardness of brass with grain size. Results of this kind are very surprising if taken at their face value, for if the width of the grain boundary region is as small as we have supposed, we should expect the mechanical properties of polycrystals to be independent of grain size as long as the size is small compared with the dimensions of the specimen and of the device used in measuring the hardness. A possible explanation of this discrepancy lies in the fact that specimens must be treated in different manners in order to obtain different grain sizes and it is possible that this treatment has a profound effect upon the mechanical properties, particularly if the materials are not perfectly pure. Good evidence for this viewpoint has been given by Corson[13] from a study of very pure copper. He found that the tensile strength and hardness of coarse-grained ingots of very pure copper are practically the same as the corresponding properties of fine-grained specimens and concluded that the weakness ordinarily observed in large-grained materials is to be attributed to the influence of gaseous impurities. His results suggest that the gaseous impurities congregate at the grain boundaries and, for reasons not yet completely understood, weaken the bond between grains. In coarse-grained specimens the total amount of grain boundary area is small compared with that in fine-grained materials so that the impurity is more concentrated and has a larger effect.

It should be noted at this point that the constraints provided by the intergranular material possibly are not the only factors contributing to variations of stress within polycrystalline materials on a scale of dimensions of the order of grain size. In this connection, Barrett and Levenson[14] have shown that the deformation within individual grains in compressed polycrystalline aluminum and iron is not uniform. In particular, the relative orientation of different parts of the grains changes progressively as compression proceeds. For example, in the case of aluminum the spread of orientation ranged from 7° to 10° for 10 percent compression, from 15° to 25° for 30 percent compression and from 35° to 45° for 60 percent compression. This variation in orientation was accompanied by the appearance of narrow bands on either side of which the orientation was different. These *deformation bands* were also observed when single crystals of aluminum were compressed by 50 percent between carefully lubricated plates. In these cases the difference in orientation on either side of the band was of the order of 3°. The origin of the bands has not yet been given a

[9] Taken from C. F. Elam's book, *The Distortion of Metal Crystals* (Oxford University Press, 1935), p. 53.

[10] R. Karnop and G. Sachs, Zeits. f. Physik **41**, 116 (1927).

[11] The Brinell hardness is determined by measuring the size of the indentation produced when a 1-mm ball is pressed into the specimen with a specified load. (See *A. S. M. Handbook* (1939), p. 112 *et seq.*)

[12] Investigations bearing on this point have been carried out by the following: G. Masing and M. Polanyi, Zeits. f. Physik **28**, 169 (1924); W. H. Basset and C. H. Davis, Trans. A. I. M. E. **60**, 428 (1919); Wood, Phil. Mag. **10**, 1073 (1930). See reference 9, p. 52.

[13] M. G. Corson, Trans. A. I. M. E. **128**, 398 (1938).

[14] C. S. Barrett and L. H. Levenson, Trans. A. I. M. E. **137**, 112 (1940).

satisfactory explanation. It seems possible, however, that they are regions such as the line *BD* in Fig. 51 at which bending takes place. If so, their appearance indicates that the stresses are not uniform throughout individual grains. This would not be surprising if the bands were observed only in the polycrystalline specimens for the non-uniformity could then be interpreted as being related to the influence of different faces of the grain. However, the observation of bands in single crystals indicates that the grains themselves are inhomogeneous on a scale of distance much larger than that of the spacing between slip bands.

8. Slip in Polycrystals

It is a natural scientific desire to hope to base a general discussion of the behavior of polycrystalline solids in various deforming processes, such as tension, compression, rolling and torsion on a few primary empirical laws. Present knowledge of the basic principles of slip in grains was outlined in the early sections of this series of articles. Although this knowledge cannot be said to be complete, inasmuch as the role of such factors as deformation bands is not yet understood, we at least have a qualitative formulation of the laws of plastic flow in single crystals. Unfortunately, as we saw in the preceding section, present knowledge of grain-boundary influence is not sufficient to complement the work on single crystals and provide us with a good foundation for treating polycrystals. In spite of this, we shall outline briefly the principal experimental facts concerning slip in polycrystals.

a. Crystalline orientation[15]

When polycrystalline specimens in which the grains are randomly oriented are subject to uniform but directional deformation several important and probably closely related changes occur. In the first place, the orientations of the grains become altered. This reorientation frequently occurs in such a way that the resultant distribution of grains is no longer random, but exhibits preferred arrangements. The degree of preferred orientation is never nearly as sharp as in a single crystal and usually depends both upon the material of which the specimen is made and upon the method used to produce the deformation. For example, when face-centered cubic metals are placed in tension, there is a tendency for both the (111) and (100) directions to become oriented parallel to the direction of tension. In aluminum the existing evidence indicates that most of the grains become oriented so that the (111) direction is along the axis of tension, but both orientations occur in other metals having the same structure. Similarly, in compression the (110) direction of face-centered cubic crystals tends to become aligned normal to the plane of compression. In some cases the final orientations of the grains in polycrystals appear to be similar to those occurring in single crystals that have been subject to the same deformation, but more often they are not. This fact indicates once again the importance of the grain boundary influence in determining the actual stresses within the grains. For example, in tension tests with single crystals of aluminum, the (110) direction tends to become parallel to the axis of tension, whereas the (111) direction tends to become aligned in this direction in polycrystals.

The reorientations occurring during rolling are of particular practical importance since much

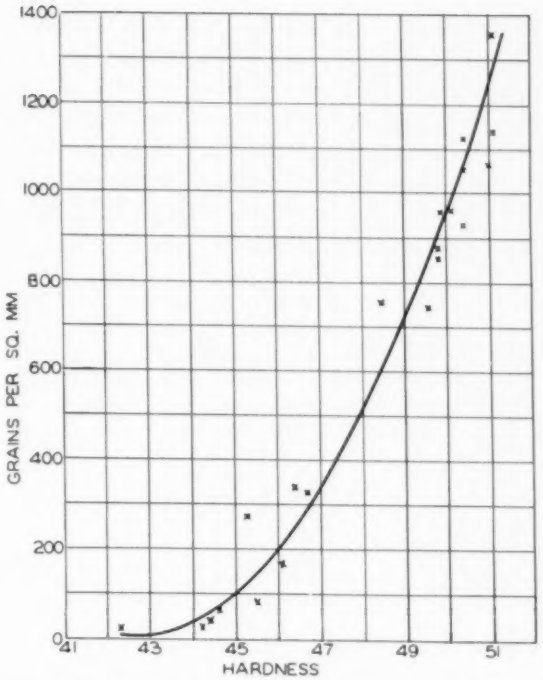


Fig. 54. Variation of Brinell hardness of brass with grain size. (After Elam.)

[15] This topic is discussed in reference 9, Chapter V.

useful metal is fabricated by this procedure. The stresses occurring in this process are more complicated than in simple tension or compression tests, because the surface of the rolled specimen is subject to normal stresses, and to shearing stresses both in the direction of rolling and at right angles to this direction. Although strong preferred orientations ordinarily are produced by rolling, the results are greatly dependent upon the conditions of rolling and cannot be summarized in terms of a few concise laws which could be conveniently discussed here. Apparently rolled orientations can sometimes be duplicated in compression tests in which the specimen is constrained so as to spread in only one direction in the plane of compression.

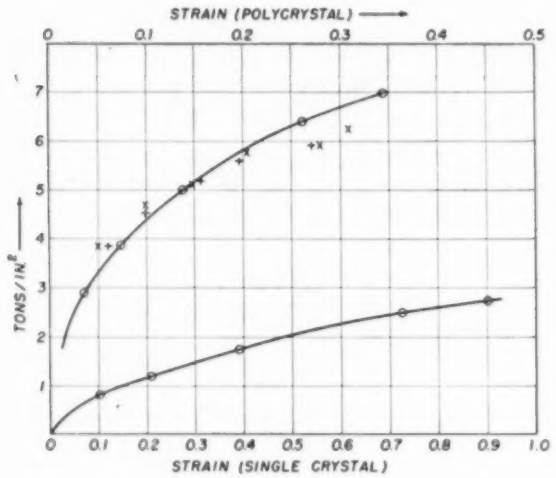


FIG. 55. The upper continuous curve is the computed stress-strain relation found by Taylor for aluminum using single crystal data. The crosses and plus signs represent experimental values. The lower curve is a representative stress-strain curve of a single crystal of the same material.

One of the interesting effects of extensive deformation on polycrystalline materials is the development of fiber structure. This structure results from the elongation of individual grains in the direction of the deformation, which gives them the appearance of fibers of woven fabrics. During the elongation neighboring grains appear to maintain close adhesion at their boundaries, for the rupture strength of the materials is not reduced except in extreme cases of working. This fact indicates once again the great strength of the intergranular bond. Moreover, it shows that such elongation takes place under great restrictions. It can be shown that the development of fiber structure would not be possible in hexagonal crystals if slip in the basal plane were the only possible means of producing deformation. Actually twinning seems to provide the additional degree of freedom in such crystals. It is also possible that inner-crystalline deformation of the type observed by Barrett and Levenson is an important means of deformation in many materials.

b. Work hardening

Polycrystals harden with cold work, as may be seen from the stress-strain curves shown in Figs. 52 and 53. Undoubtedly work hardening of the type observed in single crystals is a very significant factor in polycrystals for the latter, like the former, may be softened by heating. It is possible, however, that additional factors related to the presence of grain boundaries enter into the hardening of polycrystals. For example, there are indications[16] that complete recovery in polycrystals requires recrystallization of the specimen. In contrast with this, we have seen in Section 2 Part g that single crystals can recover completely far below their recrystallization temperature. It is not unlikely, of course, that the temperature accompanying recrystallization and not the change in structure is the important factor in determining softening. Since polycrystals recrystallize at far lower temperatures than single crystals, it is plausible to expect that polycrystals happen to recrystallize in the temperature range effective for softening, which is the same for single crystals and polycrystals.

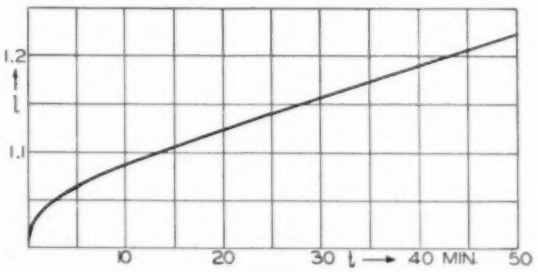


FIG. 56. Creep curve for polycrystalline cadmium. (After Andrade and Chalmers.)

A highly interesting treatment of the stress-strain curve for polycrystalline aluminum has

[16] See reference 9, p. 161.

been given by Taylor.[17] Briefly, he computed the stress-strain curve on the basis of the following assumptions:

(a) The strain is uniform in all grains of the material. This evidently contains implicitly the assumption that no slip occurs at grain boundaries, but it is obviously far more stringent.

(b) The deformation takes place in each grain by slip on the octahedral planes, as he and Elam had found for single crystals, and in such a way that the energy required is a minimum. The second part of this assumption is essentially a statement of the principle of least work.

On the basis of these assumptions, Taylor was able to compute the stress-strain curve for a polycrystal of randomly oriented grains with the use of curves for single crystals. The upper continuous curve of Fig. 55 is the theoretically derived one and is accompanied by experimental points. It may be seen that the computed curve agrees substantially with these points.

Unfortunately, the study carried on by Barrett and Levenson[14] on the orientation of grains in "homogeneously" compressed polycrystalline specimens of aluminum indicates that assumption (a) is not correct and thus detracts from the significance of the agreement found by Taylor. Moreover, Barrett and Levenson found that the average change in orientation of the grains during compression was frequently different from that predicted on the basis of Taylor's treatment and thus opens assumption (b) to question. The constructive implications of Barrett and Levenson's work have not yet been determined, but it seems definite that Taylor's theory must be modified in some essential respects, at least in the case of compression.

c. Impurity hardening

The hardening influence of soluble impurities in polycrystals seems to be qualitatively the same as in single crystals and indicates that the hardening is basically intragranular.

9. Twinning

Although twinning has been extensively investigated in polycrystals, none of this work seems to have been carried out with the view of finding differences between the behavior of polycrystals and single crystals.

[17] G. I. Taylor, J. Inst. Metals **62**, 307 (1938).

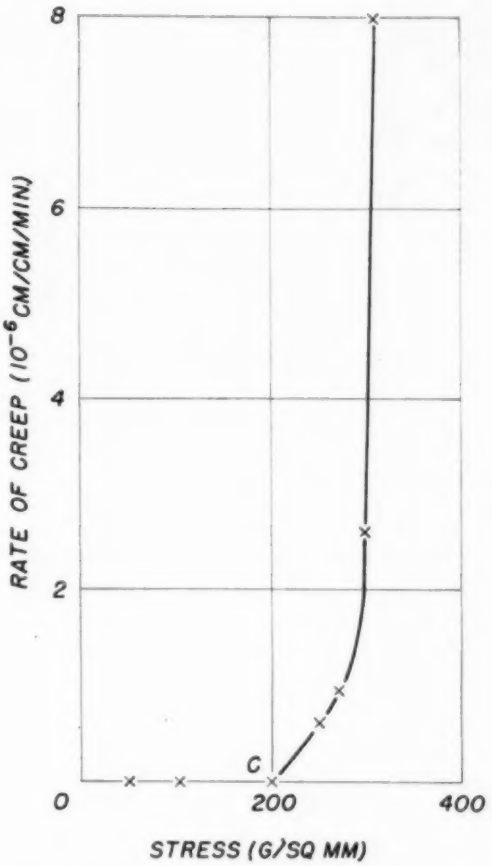


FIG. 57. Dependence of the creep rate upon stress for a bicrystal of tin. (After Chalmers.)

10. Creep[18]

Just as in the case of single crystals, the elongation *versus* time curves for polycrystals may be divided into a transient and a steady-state part. Such a curve is shown in Fig. 56 for polycrystalline cadmium[19] and the possibility of a division of this type is clearly indicated. The dependence of the steady-state creep rate upon both temperature and stress has been investigated by a number of workers[20] and seems to be

[18] For helpful discussions of this subject we are indebted to Drs. W. Kauzmann of the Westinghouse Research Laboratories and A. Lawson of the University of Pennsylvania.

[19] E. N. da C. Andrade and B. Chalmers, Proc. Roy. Soc. **A138**, 348 (1932).

[20] An analysis of available experimental data is given by W. Kauzmann, Trans. A. I. M. E. (February meeting 1941), Tech. Pub. No. 1301.

governed by a function of the form

$$R = A \exp(-(\epsilon' - b\sigma)/kT), \qquad (1)$$

where A and ϵ' are constants, σ is the applied stress and b is dependent upon temperature, but not upon stress. It is to be emphasized that this law is derived from a study of results corresponding to the practical range of stress, which is of the order of magnitude of 1 kg/mm². This is from ten to a hundred times larger than the range of stress employed in the creep experiments on single crystals of tin described in Section 3.

Figure 57 shows the creep rate as a function of stress for a bicrystal of tin, as determined by Chalmers.[21] This result indicates that Eq. (1) fails in this case for stresses of the order of 100 g/mm². Figure 57 should be compared with Fig. 44 for single crystals, for it indicates that there is a distinct difference in the creep behavior of single crystals and polycrystals at low stresses. In particular, there seems to be a limiting stress for creep in the case of the bicrystal whereas there is none in the case of single crystals.

Extension *versus* time curves for polycrystalline specimens of tin stressed in the range of Fig. 57 are shown[22] in Fig. 58. It may be seen that the steady-state rate is practically zero for stresses below 200 g/mm² and becomes finite in the range from 200 to 300 g/mm². A similar effect has been observed by Chalmers[23] in polycrystalline lead but it has not been studied in other metals.

Values of the constants appearing in Eq. (1) as determined by curve fitting of experimental results are given in Table IX. The values of b are found to vary with temperature in the manner $B \exp(aT)$ in which B and a are constants; the range of b shown in the table gives the variations over a several hundred degree temperature range.

On general grounds we might expect two sources of creep in polycrystalline materials: First, a creep analogous to that occurring in single crystals and arising from relative motion within grains and second, creep at grain boundaries. Slip corresponding to the second type of motion does not seem to be observed near room temperature in any of the metals that have been studied up to the present time. However, if it did exist we should expect to observe it just in the range in which creep ordinarily is measured. If this type of creep did occur, it presumably could be separated from inner-granular creep by investigating the dependence of creep upon grain size, for we should expect it to play a larger role in fine-grained specimens having a comparatively large amount of grain-boundary surface than in coarse-grained specimens. Actually there does not seem to be any conclusive evidence indicating that there are two contributions to creep. It will be seen in Section 12 that there is indirect evidence that intergranular flow occurs in zinc in the range from room temperature to 100°C; however, in lieu of more definite evidence, we shall discuss creep on the basis of dislocation theory.

We saw in Section 3 that two simple mechanisms of creep may be considered in discussing creep on the basis of dislocations. First, we may consider creep as caused by the relatively slow migration of dislocations, in which case

$$dS/dt = \bar{v}\lambda N, \qquad (2)$$

where N is the density of moving dislocation lines, $\bar{v}$ is their average velocity of motion and λ is the slip distance associated with the passage

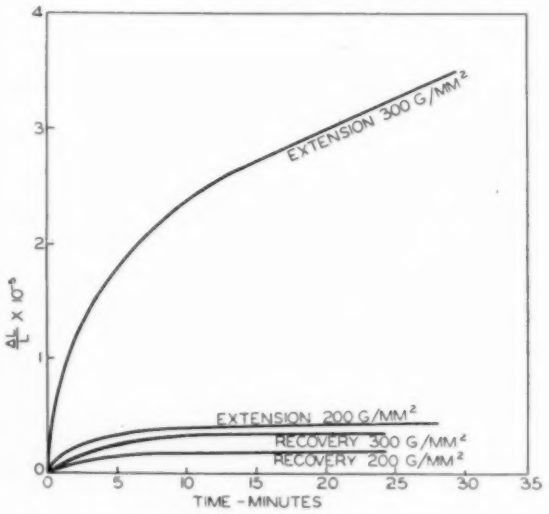


FIG. 58. Upper two curves: extension *versus* time curves for specimens of tin stressed at 300 g/mm² and 200 g/mm². The lower two curves give the subsequent contraction when the stress is released. (See Section 13.) (After Chalmers.)

[21] B. Chalmers, Proc. Roy. Soc. **A156**, 427 (1936).
[22] B. Chalmers, J. Inst. Metals **61**, 103 (1937).
[23] B. Chalmers, Proc. Phys. Soc. **47**, 352 (1935).

of one dislocation. In the steady state, N is a temperature-dependent constant so that the temperature dependence of creep rate is determined by the dependence of both v and N.

It was also pointed out in Section 3 that there is evidence that the activation energy for motion of dislocations is small in a well-annealed specimen, for which N is small. The same conclusion cannot be expected to hold during a practical creep test, since N may then be much larger. Thus it is difficult to analyze the constants appearing in Eq. (1) in terms of the quantities appearing in (2). However, we shall attempt such an analysis in the following way.

We shall assume that the rate at which dislocations are generated is given by an equation

TABLE IX. *Empirically determined creep data for metals. (After Kauzmann.) (The values of ϵ' are given in ev, those of A in units of sec.$^{-1}$ for specimens of unit length. b is given in units of A^3.*

	ϵ'	LOG$_{10}$ A	b
Pb	0.17	−8	-5×10^4
	0.1−0.35	−7 to −5	10^3-10^4
Su	0.23	−4	$5.10^3-5\times10^4$
Brass (60 Cu 40 Zn)	0.6	−4	$10^3-5\times10^4$
Steel (0.4% C)	1.2	−3	10^3-10^4

of the form

$$R_g=(N_g/\tau_g)\exp\left(-(\epsilon_g-a_g\sigma)/kT\right), \qquad (3)$$

where N_g is the number of generating centers, and that the rate of annihilation is

$$R_a=(N/\tau_a)\exp\left(-(\epsilon_a-a_a\sigma)/kT\right), \qquad (4)$$

where N is the number of dislocations present. Then in the equilibrium state

$$N=N_g(\tau_a/\tau_g)\exp\left(-(\Delta\epsilon-\Delta a\sigma)/kT\right), \qquad (5)$$

where

$$\Delta\epsilon=\epsilon_g-\epsilon_a, \quad \Delta a=a_g-a_a. \qquad (6)$$

In addition, we shall assume that

$$\bar{v}=\frac{d}{\tau_d}e^{-\epsilon_d/kT}e^{a_d\sigma/kT}, \qquad (7)$$

where d is the distance between neighboring equilibrium positions for the dislocation, τ_d is the relaxation time for passage of a dislocation from one equilibrium position to another, and

TABLE X. *Tensile strengths of annealed pure polycrystals. (The values are given in units of kg/mm².)*

Metal	Tensile Strength Obs.	Calc.	Metal	Tensile Strength Obs.	Calc.
Al	9.2	372	Pb	1.1	261
Au	12.0	782	Pd	14.	
Ca	6.4		Pt	12.	125
Cu	22.6	885	Sn	1.4	357
Mg	9.2		Zn	13.	580
Ni	32.				

the exponential coefficients give, respectively, the influence of temperature and stress on the probability that the dislocation will jump. For small stresses, the second function presumably should be a hyperbolic sine of the same argument, but we shall consider the practical range of stresses. Substituting (5) and (7) into (2), we obtain

$$\frac{dS}{dt}=\frac{\lambda d}{\tau_d}N_g\frac{\tau_a}{\tau_g}\exp\left(-(\epsilon'-b\sigma)/kT\right), \qquad (8)$$

where

$$\epsilon'=\epsilon_g-\epsilon_a+\epsilon_d, \quad b=a_g-a_a+a_d. \qquad (9)$$

Now τ_a and τ_g presumably are of the same order of magnitude, since they are associated with inverse processes. Moreover, λd is of the order of magnitude 10^{-15} cm^2, so that we obtain by comparison of (1) and (8)

$$(N_g/\tau_d)10^{-15}=A. \qquad (10)$$

We are at loss for a precise estimate of N_g, but it seems reasonable to assume that this quantity is of the same order of magnitude as the number of blocks per unit cross section, namely 10^8. The values of τ_d required to account for the observed values of A vary between 10 and 10^{-4} sec. for all of the cases listed in Table IX. The larger of these values of τ_d agrees in order of magnitude with that obtained in Section 3 from a similar analysis of the rate of creep of single crystals of tin on the basis of Eq. (2). We note again that if Eq. (2) is correct, it seems to imply that the natural relaxation time for motion of dislocations is very long compared with an atomic oscillational period.

A second possible simple mechanism for creep is based on the notion that dislocations move very rapidly a distance L from the point at

which they are generated and then become stuck. The corresponding equation was discussed in Section 3 and is

$$dS/dt = \lambda L dN/dt. \qquad (11)$$

As we have seen in Section 3, the purely transient creep observed in single crystals of tin, zinc and cadmium at low stresses can be explained more suitably on the basis of Eq. (2) than on the basis of (11), but it is possible that Eq. (11) is valid in the practical range of creep.

Substituting into (10) from Eq. (3), we obtain

$$\frac{dS}{dt} = \lambda L \frac{N_g}{\tau_g} \exp\left(-(\epsilon_g - a_g\sigma)/kT\right). \qquad (12)$$

Thus in the present case

$$A = 10^{-8} L N_g/\tau_g.$$

If L is of the order of magnitude 10^{-4} cm, as is assumed in Taylor's theory of hardening, we obtain

$$A = 10^{-12} N_g/\tau_g.$$

In the cases listed in Table IX, N_g/τ_g varies between 10^9 and 10^4. If our previous estimate of N_g, namely 10^8, is valid, we obtain the result that τ_g varies between 10^{-1} and 10^4 sec., which seems to be much too large to be reasonable. In other words, the parameter values derived from the use of Eq. (2) seem to be more reasonable than those derived from (11). It must be admitted, however, that the evidence in favor of Eq. (2) is far from satisfying. At most, we may say that the theory of dislocations seems to provide a rough semi-quantitative means of correlating creep data.

Comparing Eqs. (1) and (6), we see that the quantity b in (1) is composite if Eq. (2) is correct. It is interesting to note that the experimental values of b are so large that $b\sigma$ is of the order of 1 ev for stresses near 1 kg/mm². This implies that either a_g or a_d in Eq. (9) may be large. In the first case, it would be concluded that there is very large stress magnification near the slip nuclei and in the second that dislocations have lengths of the order of 10^{-5} cm. It is not easy, however, to understand the strong dependence of the experimental values of b upon temperature. Kauzmann[20] has suggested that this dependence is related to an increase with temperature of the lengths of either slip nuclei or dislocations. Additional evidence for such a variation possibly is furnished by the difference between the values of τ we have obtained above from comparisons of Eqs. (2) and (11) with Eq. (1) and the value we might expect to obtain, namely 10^{-13} sec. According to Kauzmann's suggestion the increase in the length of dislocations with temperature gives rise to an increase of the activation energy with temperature. If the creep data were analyzed on the basis of Eq. (1), taking into account a temperature dependence of ϵ', a value of A larger by a factor of 10^{10} could be obtained. In this case τ would come out to be of the order of magnitude of 10^{-10} sec.

11. Rupture in Polycrystals

It was pointed out in Section 7 that rupture is transcrystalline except at temperatures just below the melting point. This indicates that grain boundaries are not sources of great weakness at sufficiently low temperatures and we might expect to correlate the breaking strengths of polycrystals with those of single crystals.

Such a correlation is made difficult for several reasons. In the first place, stresses within grains are not the same as the applied stress because of grain boundary influence. In the second place, specimens of the same material prepared in different ways will extend differently during a rupture test and hence will usually have different orientations when fracture occurs. Fortunately, the second factor may be eliminated to an appreciable extent by comparing specimens which already possess the maximum degree of preferred orientation of the type that would be produced during the rupture test.

Now we saw in Section 5 that each crystallographic plane of a single crystal appears to possess a characteristic tension stress at which it will rupture. Moreover, evidence was cited which indicates that this stress is independent of previous deformation. Reasoning from this we should expect the rupture stress of a polycrystalline metal to be independent of deformation as long as the deformation does not produce widely different orientations. This expectation seems to be borne out in the case of copper. Figure 59

shows[24] stress-strain curves for specimens of copper wire prepared from the same material. Prior to the test the wires were drawn to different diameters, during which process comparable degrees of preferred orientation presumably were attained. The specimens were then subjected to tension stress and the strain was measured in terms of the contraction in cross section. The specimens began to neck at the stress corresponding to point 2, and fracture occurred at the points 3. It may be seen that the breaking

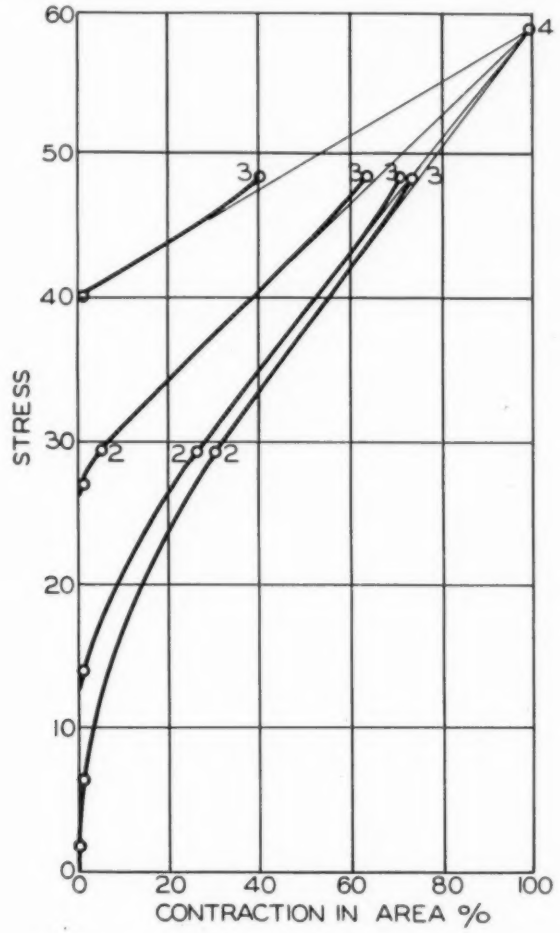


FIG. 59. Stress-strain curves for variously worked specimens of polycrystalline copper. It may be seen that the rupture stresses are closely the same in each of the cases. The strain is measured in terms of the contraction in cross-sectional area.

stresses are nearly the same in all four cases, which indicates that the previous strain had little effect on the breaking stress.

[24] J. v. Mollendorf and J. Czochralski, Zeits. Verein. deuts. Ing. **54**, 931 (1913).

Just as in the case of slip, there seems to be evidence that the grain size has an important influence on the breaking strength even when the crystal size is small compared with the dimensions of the specimen. As was pointed out in Section 7, this result is inexplicable unless it is assumed that the different treatment required to produce various grain sizes results in different distribution of impurities at grain boundaries. This interpretation of the effect is supported by the experiments of Corson, discussed in Section 7.

A comparison of experimental breaking strengths and those computed from Polanyi's equation with the use of experimental values of the surface tension is given in Table X. It may be seen that the experimental values are smaller than the theoretical ones by a factor of the order of a hundred or more, just as in the case of single crystals. There is a slight indication that polycrystals are somewhat stronger than single crystals, but this effect may be related to the redistribution of stresses resulting from the influence of grain boundary restrictions.

12. INTERNAL FRICTION

The principal source of internal friction in nonferromagnetic single crystals seems to be intimately associated with plasticity, as we have seen on Part *d* of Section 2. The same source is sometimes important in polycrystals, but, as was first pointed out by Zener,[25] is often completely masked by another. The second source has been extensively studied by Zener and his co-workers and arises from the thermoelastic effect in the following way.

When a crystal is stressed suddenly its temperature changes, heat being generated or absorbed, depending on the sign of the dilatation. The change in temperature is proportional to the stress and reverses its sign if the sign of the stress is reversed. Since the stresses in a polycrystal are not uniform but vary from grain to grain, it follows that the temperature resulting from the thermoelastic effect will vary from point to point. Now if the stresses vary slowly compared with the time required for heat to flow

[25] C. Zener, Phys. Rev. **52**, 230 (1937); **53**, 90 (1938); W. Otis and R. Nuckolls, Phys. Rev. **53**, 100 (1938); R. H. Randall, F. C. Rose and C. Zener, Phys. Rev. **56**, 343 (1939); C. Zener, Proc. Phys. Soc. **52**, 152 (1940).

between regions of different temperature, the temperature will remain constant and the transfer of mechanical energy into heat, and vice versa, takes place reversibly. In this case the net amount of heat generated during a complete cycle is zero and there is no contribution to the internal friction. Similarly, if the variations in stress take place so rapidly that no heat flows during a stress cycle, the process is said to occur adiabatically since each region of the material behaves as if it were thermally isolated. It follows that the net conversion of elastic energy into heat during a complete cycle also is zero in this case. However, if the frequency or the thermal conductivity are such that the process is neither isothermal nor adiabatic, the conversion will not be reversible and there will be a finite contribution to the internal friction from the thermoelastic effect.

Zener has pointed out that if the average applied stress is uniform throughout the specimen, the critical distances for thermal diffusion in a polycrystal should be the dimensions of the grain, for stresses should vary relatively abruptly from one grain to another. Reasoning from this on the basis of dimensional analysis, he suggested that the internal friction should depend on the dimensionless quantity $\nu d^2/D$. Here d is the grain diameter, ν is the frequency and D is the thermal diffusion constant defined by the equation

$$D=\frac{\text{thermal conductivity}}{(\text{specific heat})(\text{density})}.$$

In other words, for a given material the relative value of the internal friction, should be a universal function $g(\nu d^2/D)$. Figure 60 shows schematically the manner in which this function should depend on its argument. It has a peak near the value unity and decreases to zero on either side. This type of dependence of internal friction has been verified in brass.

If the average stress is not uniform throughout a specimen, as during the transverse vibration of a reed, there will be an additional contribution to the thermoelastic internal friction resulting from the variations in temperature across the specimen. This contribution is determined by a function similar to that shown in Fig. 60, the argument being $\nu L^2/D$ in this case, where L is the distance between the points of maximum temperature difference, which is the thickness of the specimen in the case of the vibrating reed. The thermoelastic internal friction was first discovered and studied systematically in inhomogeneously stressed systems of this type.

The factors contributing to the absolute value of the thermoelastic internal friction have been investigated by Zener, but will not be discussed here since they would lead us too far afield. This work shows that in general we should expect the internal friction arising from stress variations between grains to be smaller the more nearly elastically isotropic the material is.

We might expect to find peaks of the type shown in Fig. 60 in single crystals as a result of stress variations either from domain to domain or between larger regions. The first of these cases corresponds to an effective grain size of 10^{-4} cm and would produce a peak in the megacycle range of frequency. This type of internal friction actually seems to be negligible in Read's experiments.

It should be mentioned that the value of D will be dependent on the ease with which heat

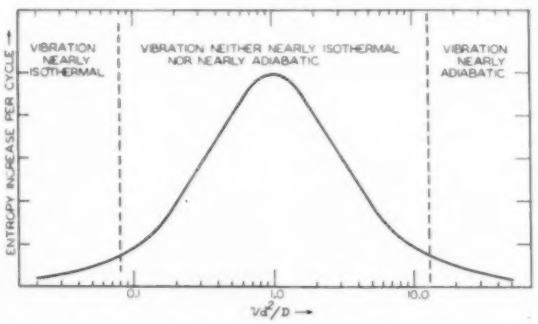


FIG. 60. Relative value of the internal friction as a function of the dimensionless variable $\nu d^2/D$. (After Zener.) The function has its peak in the region where the argument is unity and decreases on either side.

flows across grain boundaries and is not necessarily an average of the values for single crystals. Actually, the influence of grain boundaries has not been studied.

The separation of the internal friction of plastic origin from the thermoelastic contribution can be carried out in several ways. For example, if measurements are carried out for frequencies extending into the adiabatic or isothermal regions, the residual value of plastic origin will appear as

an additive constant on a function of the form of Fig. 60. The range of frequency needed for this evidently will depend upon d and D. Thus in the case of three specimens of brass of different grain size,[26] a residual decrement of $0.15.10^{-5}$ was determined by extrapolating to the region of low frequencies.

Barnes and Zener[27] carried out a similar separation at fixed frequency in zinc (impurities 0.5 percent). Measurements were made on specimens of four grain sizes for various temperatures in the range from 0°C to 100°C. The values of $\nu d^2/D$ at the frequency used were 0.008, 0.008, 0.02 and 0.8 for the four specimens. In the first three cases the internal friction varied with temperature in the manner $A \exp(-\epsilon/kT)$, the constant ϵ being the same in all cases and the constant A varying inversely as the square of the grain size. In the fourth case, in which the thermoelastic internal friction should be near its maximum value, the internal friction was expressible in the form

$$\Delta = C + A_4 \exp(-\epsilon/kT) \qquad (1)$$

in which C was independent of temperature and equal to $44 \cdot 10^{-5}$, ϵ was the same as in the other three cases, and the ratio of A_4 to the values of A occurring in the other cases varied inversely as the grain sizes. The investigators conclude that the temperature-dependent term had its origin in the plasticity, whereas the constant in (1) corresponds to the thermoelastic dissipation. They also conclude from the dependence of the coefficient of the Boltzmann factor upon grain size that the source of the temperature-dependent internal friction lies on the grain boundaries, rather than in the interior of the grains.

13. Secondary Plastic Effects

In addition to the primary plastic effects, slip, creep, twinning and rupture, there are a number of interesting secondary effects which, like fatigue, undoubtedly find their origin in the primary effects. Some of these are observed only in polycrystals; others are observed in single crystals as well. We shall discuss both here.

[26] C. Zener and H. Randall, Trans. A. I. M. E. **137**, 41 (1940).

[27] A. H. Barnes and C. H. Zener, Phys. Rev. **58**, 87 (1940).

a. Creep recovery

If a specimen of polycrystalline material is extended by creep, it is found to contract slowly with continually decreasing rate when the applied stresses are removed. The lower two curves in Fig. 58 show this effect for two specimens of polycrystalline tin. In these cases the ordinate gives the decrease in length as a function of time after the specimens which yielded the upper two curves were unloaded. The rate of contraction is rapid at first and then decreases to zero. This effect did not occur in the single crystals studied by Chalmers, which indicates that it is primarily a polycrystalline one. The effect is also observed after practical creep tests.

This type of creep recovery apparently can be given a satisfactory explanation on the basis of a principle first proposed by Masing.[28] He pointed out that the type of extension occurring in different grains of a polycrystal during any deformation will be different because of differences in orientation. Some grains will deform almost entirely as a result of slip whereas others will deform elastically. For small deformations, such as those dealt with in Chalmers' experiments, we may expect a continuous range between grains which are deformed almost entirely elastically and those which have deformed almost entirely as a result of plastic flow. Consequently the actual stress in the interior of the metal will vary from a maximum value in those grains which have undergone the least plastic deformation to a minimum value in those which have undergone the most. Thus when the externally applied load is removed and the average stress reduced to zero, some of the grains will be stressed in tension and some in compression, the latter being those which had previously flowed. If the temperature is sufficiently high, these stresses will cause creep and the specimen will be observed to contract in the manner of Fig. 58.

b. The Bauschinger effect

If the previous interpretation of creep recovery is correct, we might expect that just after a polycrystalline specimen has been unloaded, following an elongation or compression, there are forces present in some of the grains which would

[28] G. Masing, Wiss. Siemens Konzern **3**, 231 (1924); **4**, 74, 244 (1925); **5**, 135, 142 (1926).

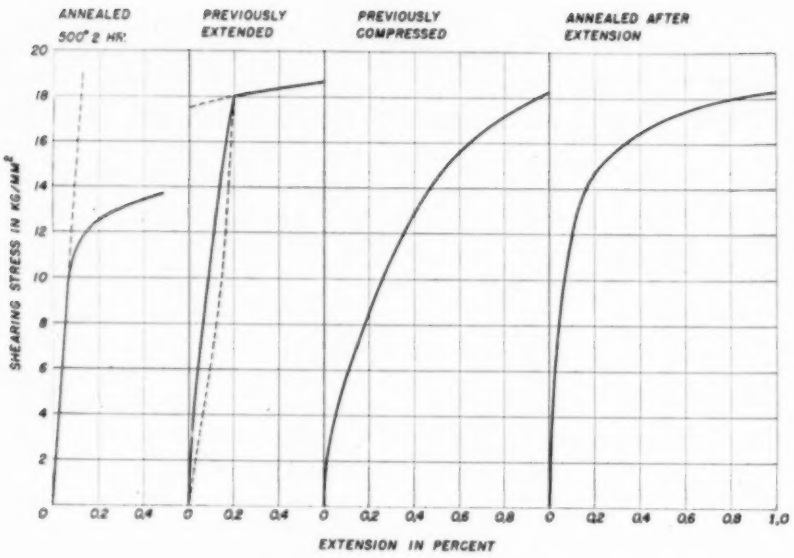


FIG. 61. Illustration of the Bauschinger effect in fine-grained polycrystalline brass. (a) shows the extension *versus* stress curve for the virgin material. (b) is the stress-extension diagram obtained after previous extension. (c) is the same curve obtained after previous *compression*. It may be seen that the elastic range is very small, illustrating the Bauschinger effect. (d) is the stress-strain curve similar to (c) obtained when the specimen is annealed at relatively low temperatures after compression. The hardness remains but the Bauschinger effect is gone.

aid in producing a deformation of opposite sign. Thus we might expect the critical stress required to produce plastic compression following a small elongation to be lower than the critical compression stress in the virgin state. An effect of this kind was first discovered in polycrystals by Bauschinger[29] and was explained qualitatively in the basis of Masing's principle, discussed in part *a* of this section. Figure 61 shows[30] the Bauschinger effect in fine-grained brass. Figure 61(a) is the stress-strain curve obtained during an extension of a virgin specimen of the material by a few percent. The critical shearing stress is about 10 kg/mm². Essentially the same value would have been obtained if the virgin specimen were compressed rather than extended. Figure 61(b) shows the stress-strain curve for compression following the extension. It may be seen that the critical shearing stress for this deformation is very small for the range of compression corresponding to the original extension. Figure 61(c) shows the type of stress-strain curve obtained if the specimen is further extended (without compression) after the original extension, and exhibits the usual effects of work hardening. The last figure shows the stress-strain curve obtained in compression after annealing the originally extended specimen at 150° for several hours. The Bauschinger effect has disappeared, but the hardness remains, showing that the factors contributing to the Bauschinger effect are not immediately connected with the hardening.

A quantitative development of Masing's theory of the Bauschinger effect was carried out by Heyn,[31] who showed that curves of the type of Fig. 61(b) could be obtained for polycrystals in a reasonable manner on the basis of the theory.

It is unfortunate for the simple theory that a well-defined Bauschinger effect has been observed by Sachs and Shoji[30] in single crystals of brass. A typical example of their results is shown in Fig. 62. In this the ordinate is stress and the abscissa is strain, the positive and negative directions corresponding, respectively, to tension and compression. The curve obtained in a stress-strain cycle starting with a virgin specimen is somewhat reminiscent of a magnetic hysteresis curve. The segment D_1 of the entire curve represents the stress-strain curve obtained in the first extension and rises sharply in the elastic region, as in Fig. 61(a). When the tensile stress was relieved, the elastic contraction to the point A occurred. Following this, the specimen was compressed and the curve D_2 was obtained. It is evident that the decrease from the point A is much more gradual than the original rise from the origin, showing the Bauschinger effect. If this part of the curve were inverted about the point

[29] J. Bauschinger, Ziviling. **27**, 289 (1881).
[30] G. Sachs and H. Shoji, Zeits. f. Physik **45**, 776 (1927).
[31] E. Heyn, *Festband Kaiser-Wilhelm Gesellschaft* (1921), 131.

A, it would bear the same relation to D_1 that Fig. 61(b) bears to Fig. 61(a). The main difference between the polycrystalline and single crystal specimens appears to lie in the fact that the latter are somewhat softer, as may be seen by comparing the ordinates in Figs. 61 and 62. The curve D_3 represents another stress-strain curve for extension following the compression D_2.

A summary of the experimental facts obtained from this type of investigation on single crystals of brass is as follows:

1. The primary stress-strain curves obtained in tension or compression for virgin materials are nearly identical.

2. The Bauschinger effect is observed in the opposite deformation following the primary one, regardless of the sign of the primary one. We shall call the deformation following the primary the secondary one. The secondary curves are very nearly identical if the primary extensions or compressions are identical.

3. Secondary curves, and curves such as D_3 obtained by later deformation, are very nearly alike as long as the over-all extension and compression remains fixed in each cycle. The principal difference lies in the fact that a gradual work hardening takes place; that is, the curves rise to larger absolute values of the abscissa.

4. If at any period in a sequence of cycles, a deformation larger than the preceding is carried out, the Bauschinger effect will be more clearly marked in the deformation immediately following.

5. Annealing for two hours in the range from 250°C to 400°C removes the Bauschinger effect for the next deformation, but does not appreciably affect the work hardening.

6. When correction is made for the slightly greater hardness of polycrystals, the Bauschinger effect in single crystals of brass may be said to be fully as well defined as in polycrystals.

At first sight, the existence of a well-developed Bauschinger effect in single crystals seems to be in contradiction with Masing's theory. Actually, there are two ways of interpreting the results:

(a) In the first place it is possible that the single crystals of brass on which the measurements were made actually are far less homogeneous than the best single crystals of monatomic substances. Thus it is possible that the deformations are not homogeneous and that parts of the crystal undergo almost purely elastic deformations. This possibility could be tested by more extensive investigations of single crystals of other materials. Apparently, pure monatomic crystals are not well suited for tests of this type because of their great softness. It should be added that Sachs and Shoji made measurements on single crystals of aluminum-copper alloys, but found the effect smaller than in brass and did not extend the investigation. At the present time, this interpretation seems to be the most plausible one.

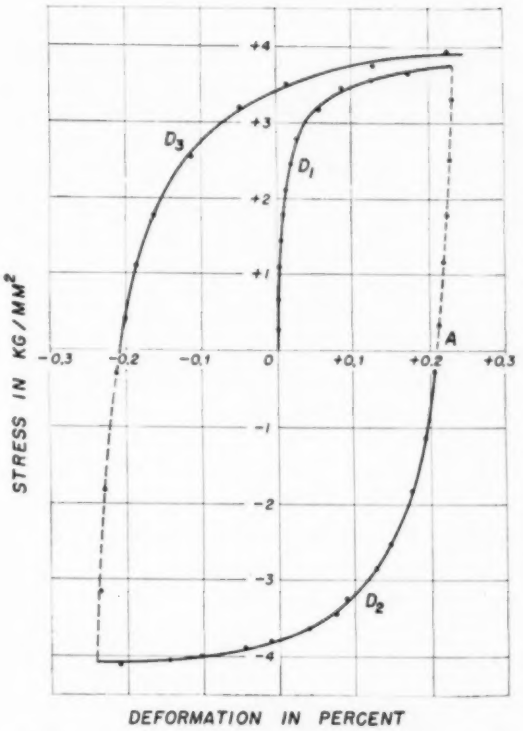


FIG. 62. Stress-strain diagram for single crystal of brass obtained during a cyclical deformation process. Negative strains correspond to compressions, positive to elongations. D_1 corresponds to the virgin material.

(b) In the second place, it is possible that there are two contributing causes for the Bauschinger effect in polycrystalline materials—one being the effect suggested by Masing, and the other an effect that occurs even in single crystals. Since Heyn's work indicates that Masing's theory is sufficient to explain at least the magnitude of the effect in polycrystals, we must conclude that if two co-existing effects occur they are of comparable importance.

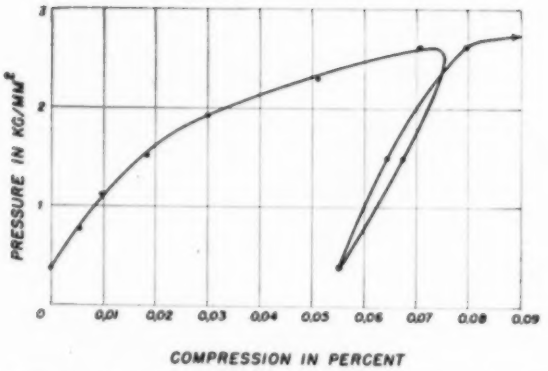


FIG. 63. The elastic after effect in a compressed single crystal of brass.

It has recently been suggested by Masing that the theory of dislocations may contain features usable for an explanation of the Bauschinger effect in single crystals. In particular, he has proposed the possibility that dislocations which have moved into the crystal as a result of the action of a given stress move backward under a lower reverse stress than is needed to make them move forward. According to Masing, this unidirectional hardening is a consequence of the fact that the newly formed dislocations will encounter dislocations of opposite sign, generated on the opposite sides of the particular block in which they are moving, if they continue to proceed in the direction in which they were started. They will not encounter these dislocations if their direction of motion is reversed.

This picture seems difficult to justify on the basis of the theory of dislocations developed here for we should either expect the positive and negative dislocations in the same block to attract and lower the shearing stress for motion in the same direction, or we should expect them to become intermixed in the manner suggested by Taylor and form a lattice that possesses equal rigidity in either direction. The following, however, is an alternate possibility. In discussing resoftening of work-hardened crystals, it was pointed out that newly formed dislocations near surfaces are subject to a strong attractive "image" force which drops off as the first power of the distance from the surface. It is readily seen that this force operates in just such a direction as to contribute to the Bauschinger effect. Evidently a quantitative estimate of the magnitude of this contribution is necessary before we can draw a definite conclusion as to whether the mechanism is a likely one.

The fact that the Bauschinger effect can be completely removed by an annealing process that does not produce complete resoftening seems to indicate strongly that dislocations are not primarily responsible for the effect.

c. The elastic after effect

Another interesting secondary effect observed commonly in polycrystals and in some single crystals, such as the brass specimens used by Sachs and Shoji, is the elastic after effect illustrated in Fig. 63. It is found that the stress-strain curve obtained when a specimen is unloaded and then reloaded, following an extension, is not a single-valued function but has the form shown in the figure.

This effect may readily be explained in a qualitative manner on the basis of Masing's principle of inhomogenous strain. We may postulate that when the specimen is first unloaded the contraction is entirely elastic. As soon as the elastic forces in the regions that have undergone the most extensive plastic deformation are released, the remaining regions will exert on them stresses of such sign as to reverse the original deformation. These stresses increase as the unloading proceeds and cause a small amount of irreversible plastic flow, analogous to the flow occurring during creep recovery (Fig. 58). In fact, it seems likely that such creep recovery actually would have been observed on the single crystal specimens of brass used in obtaining Fig. 63.

Evidently, the questions that arise concerning Masing's theory and the existence of the elastic after effect in single crystals of brass are basically similar to those discussed in connection with the Bauschinger effect in single crystals.

If we grant that all three of the secondary plastic effects presented in this section have similar origin, we may conclude from the fact that single crystals of tin do not show appreciable creep recovery that they would not show a Bauschinger effect or elastic after effect comparable with that observed in brass. From this viewpoint, then, it follows that the brass crystals on which Sachs and Shoji made their observations were not as homogeneous as the single crystals of tin studied by Chalmers.

Here and There

University of Michigan Summer Session

In addition to a number of degree courses in physics, the University of Michigan Summer Session, 1941, is presenting a Symposium in Theoretical Physics extending from June 30 to August 22. Lectures in this symposium will be given by Professor Wolfgang Pauli, Institute for Advanced Study, Princeton University; Assistant Professor Frederick Seitz, University of Pennsylvania; Dr. Julian Schwinger, University of California; and Assistant Professor Victor F. Weisskopf, University of Rochester.

★

Dayton C. Miller Research Fellowship

In memory of Dr. Dayton C. Miller, who was connected with Case School of Applied Science for more than fifty years prior to his death last February, the Dayton C. Miller Research Fellowship in Physics has been established, according to an announcement received from Robert S. Shankland, present Head of the Physics Department at Case. "We feel that this fellowship will do much to preserve the traditions of research in pure science that Dr. Miller established here," Dr. Shankland states.

Dr. Miller's will made provision for a fellowship award to be available at a later date, but funds set aside by his friends and former students have made possible its establishment this year. The first recipient of the award, to be made to a graduate assistant in physics during his second year of work towards an advanced degree, is P. J. Rice. Mr. Rice plans to continue his research studies on the electron microscope.

★

Symposium on Light and Lighting

"Light and Lighting" will be the topic of a symposium to be sponsored by the Ohio Section of the American Physical Society at the October, 1941, meeting at Cleveland, Ohio, according to a preliminary announcement made by Richard H. Howe, Secretary of the Section.

★

Members of Scientific Mission Return

Dr. Warren Weaver, Director of the Natural Sciences of the Rockefeller Foundation and Edward J. Poitras, of the California Institute of Technology, returned to New York by clipper on April 28, having spent five weeks in England as members of the scientific mission appointed by President Roosevelt from the membership of the National Defense Committee. Dr. Weaver, Mr. Poitras, Professor Louis B. Slichter, of the Massachusetts Institute of Technology, and Dean John T. Tate, of the University of Minnesota, were together in London throughout the nine-hour bombing raid of April 16 and 17—the most severe London has experienced.

Science

Flight Similitude Chamber Aids Meteorologists

Capable of simulating temperature and pressure conditions of altitudes up to thirty miles, a flight similitude chamber is being used at the University of Chicago to train radiosonde operators, urgently needed in the Government's long-range weather forecasting program. Dr. Michael Ference, Jr., instructor in physics and a member of the staff of the University of Chicago's Institute of Meteorology, designed the "stratosphere box," which is one of three in the United States.

The University of Chicago's Institute of Meteorology is headed by Dr. Carl G. A. Rossby. Dr. Rossby is largely responsible for the development of five-day forecasts through his work with "three-dimensional" airflow charts. Thirty army, navy, air corps and weather bureau students are being trained at the Institute for government service.

★

Officers Elected for the Ohio A. P. S. Section

The fifth meeting of the Ohio Section of the American Physical Society was held in joint session with the Ohio Academy of Science at Case School of Applied Science and Western Reserve University, Cleveland, Ohio, Saturday, May 10. Officers of the Ohio Section elected for the year 1941–1942 were: W. E. Forsythe, *Chairman*, General Electric Company, Nela Park; Paul B. Taylor, *Vice Chairman*, Aircraft Radio Laboratory, Wright Field; Richard H. Howe, *Secretary-Treasurer*, Denison University.

★

Radio Amateur Wins Award for Radio School

Marshall H. Ensor, high school teacher and radio amateur of Olathe, Kansas, received the fourth annual William S. Paley Amateur Radio Award recently for his outstanding service to the nation in training thousands in the radio code and radio fundamentals over a ten-year period, during which he broadcast regular lessons in those subjects from his own transmitter to thousands of "air students." Paul W. Kesten, Vice President of the Columbia Broadcasting System, awarded the trophy on behalf of its donor, William S. Paley, CBS President.

★

British Scientists to Work in America

Sir Lawrence Bragg, Cavendish Professor of Experimental Physics at the University of Cambridge, and Dr. Charles Darwin, Director of the British National Physical Laboratory, recently left England for Canada. Sir Lawrence will work in Ottawa at the National Research Buildings in cooperation with Professor Fowler as liaison officer with Canadian scientific men. Dr. Darwin will work in Washington.

Dr. Horace S. Uhler, Professor of Physics at Yale University is retiring from active service after being connected with the University for thirty-four years, according to a recent note in *Science*.

Louis E. Levy Medal Awarded

"Combined Stress Experiments on a Nickel-Chrome Molybdenum Steel," a paper published in the August, 1940, issue of the *Journal of the Franklin Institute*, has won the Louis E. Levy Medal for its authors, John M. Lessels and Charles W. MacGregor. The co-recipients of the annual award are both Associate Professors in the Department of Mechanical Engineering at the Massachusetts Institute of Technology.

The Levy Medal, presented annually by the Committee on Science and the Arts of the Franklin Institute, is awarded "to the author of a paper of especial merit, published in the *Journal of the Franklin Institute*, preference being given to one describing the author's experimental and theoretical researches in a subject of fundamental importance."

★

Bulletin of Mathematical Biophysics

The September, 1941, issue of the *Bulletin of Mathematical Biophysics*, edited by N. Rashevsky and published by the University of Chicago Press, will contain the following articles:

Electrical Charges and Potentials in Cells Resulting from Metabolism of Electrolytes, ROBERT R. WILLIAMSON

Note on the Mathematical Biophysics of Temporal Sequences of Stimuli, N. RASHEVSKY

A Note on the Nature of Correlations Between Different Characteristics of Organisms, N. RASHEVSKY

Mathematical Biophysics of the Galvanic Skin Response, CLYDE H. COOMBS

A Theory of Steady-State Activity in Nerve-Fiber Networks: II. The Simple Circuit, ALSTON S. HOUSEHOLDER

★

National Academy of Sciences Elects

Among those elected to membership by the National Academy of Sciences during its April meeting held in Washington, are the following physicists: Charles Christian Lauritsen, Professor of Physics, California Institute of Technology, Pasadena; Alfred Lee Loomis, Director, The Loomis Laboratories, Tuxedo Park, New York; J. Robert Oppenheimer, Professor of Physics, University of California, and California Institute of Technology; and Joseph Slepian, Research Engineer, Westinghouse Electric and Manufacturing Company, Pittsburgh, Pennsylvania.

This spring meeting of the Academy was also the occasion of the presentation of the Henry Draper Medal to Dr. Robert Williams Wood, Research Professor of Experimental Physics at the Johns Hopkins University, "in recognition of his contributions to astronomical physics."

★

Five Research Fellows Appointed

Five new Westinghouse Research Fellows have received appointments enabling them to carry on their studies at the Westinghouse Research Laboratories in East Pittsburgh, it was announced recently by Dr. E. U. Condon, Associate Director of the Laboratories. The appointees are: Dr. Daniel Alpert, Stanford University; Dr. John W. Coltman, University of Illinois; Dr. Thomas W. Dakin, Harvard University; Dr. Theodore Holstein, New York University; and Dr. James S. Koehler, University of Pennsylvania.

Comprising the fourth group selected under the Westinghouse Research Fellowship plan inaugurated in 1938, the appointees will carry out original research on scientific problems of their own choosing in an attempt to discover new fundamental facts in the realm of pure science.

★

Calendar of Meetings

July

23–25 American Society of Civil Engineers, San Diego, California

August

27–29 American Institute of Electrical Engineers, Yellowstone National Park

September

2–6 American Mathematical Society, Chicago, Illinois
8–12 American Chemical Society, Atlantic City, New Jersey
8–10 American Astronomical Society, Williams Bay, Wisconsin
22–26 American Association for the Advancement of Science, Chicago, Illinois
25–26 Society of Automotive Engineers, Millwaukee, Wisconsin

October

6–10 National Safety Council, Chicago, Illinois
12–15 American Society of Mechanical Engineers, Louisville, Kentucky
13–16 Society of Motion Picture Engineers, New York, New York
15–17 American Society of Civil Engineers, Chicago, Illinois
20–24 American Society for Metals, Philadelphia, Pennsylvania
20–24 American Welding Society, Philadelphia, Pennsylvania

New Book

Industrial Research Laboratories of the United States

Seventh Edition. Pp. 372. National Research Council, Washington, December, 1940 (National Research Council Bulletin No. 104). Price $3.50.

The revised edition of this directory contains information concerning 2264 industrial concerns which maintain one or more research laboratories. This list, with certain subtractions, includes 469 more concerns than were listed in the sixth edition which was issued in December, 1938.

The directory gives, in most cases, the following items for the laboratories listed: address of the laboratory; name of the president of the corporation; name of the research director; classification and number of members of the research staff; research activities; and cooperative research and laboratory publications when the company sponsors these enterprises.

There are several indexes to the directory which cover the geographical distribution of laboratories, the personnel, and the subject of the research activities. The personnel index includes the names of about 6900 research workers.

Contributed Original Research

Temperature Correction Methods in Calorimetry

ALLEN KING AND HORACE GROVER
Rensselaer Polytechnic Institute, Troy, New York

(Received March 4, 1941)

There are two fundamentally different methods of correcting calorimetric data for heat losses to surroundings. The first method, applicable only to experiments in which the supply of heat to the calorimeter can be neglected after a certain time, uses the afterperiod cooling curve, together with a similar curve from data taken before the heat supply started, to evaluate appropriate correction terms. This problem is discussed from a general point of view and correction formulas previously reported by Regnault and Pfaundler, Schultes and Nübel, Roth, and others are shown to be variations of a general scheme. A second method is needed in such experiments as the measurement of specific heat when it is not safe to neglect the heat supplied by the hot body even after a very long time. In this case, there is no afterperiod cooling curve and additional assumptions concerning the heat exchange between the calorimeter and the hot body within it must afford the needed information. The assumptions originally suggested by Rowland are re-examined. It is shown that previous investigators (Rowland and Hoare) deduced inadequate correction methods from these assumptions. A better method is developed and a simple approximation of it is suggested. Application of the new method to a set of data gives results different from those obtained by the Rowland-Hoare methods, from those obtained by the Regnault-Pfaundler method (which, despite previous use, should not be applied to such data), and from those obtained by the method of Dickinson (which is recommended by the ASTM standards). The differences are significant in precision work. A brief critical review of graphical methods of obtaining the temperature correction concludes the paper. This review includes a graphical analog of the new method mentioned above.

ALTHOUGH extraneous thermal exchanges in calorimetric experiments may be controlled, usually they cannot be neglected. For this reason, numerous methods have been devised to obtain an approximate correction from actual measurements.[1] Some of these methods are erroneous; others are unnecessarily elaborate. Often methods suitable for one type of experiment have been applied to another type for which they are inadequate. We shall consider here the general problem of correction for thermal exchange and discuss the conditions appropriate to the several cases that occur in practice.

Calorimeter data usually consist of readings of temperature as a function of time.[2] At the outset, the time rate of change of these calorimetric temperatures will be assumed to consist of two terms. One, the "heating function," depends in general upon the temperature difference between the source of the heat to be measured and the calorimeter fluid; and the other, the "cooling function," depends upon that between the calorimeter and the surrounding jacket. The source of heat may be a hot solid as in specific heat measurements, an ignited bomb as in combustion experiments, or simply a hot wire as in certain electrical equivalent of heat experiments. The calorimeter "fluid" frequently is water, but may even be copper. Following the usual practice, we shall suppose that the temperatures of the source and the fluid are uniform throughout the respective bodies.

In symbols the above assumption is represented by the expression

$$d\theta/dt = F(\Theta-\theta)+f(\theta-\theta_j) \qquad (1)$$

in which Θ is the temperature of the source of heat, θ is the temperature of the calorimeter fluid, and θ_j is the temperature of the jacket. The two functions are expanded into power series in terms of their respective arguments, so that

$$F(\Theta-\theta) = h+b(\Theta-\theta)+\cdots \qquad (2)$$

and

$$f(\theta-\theta_j) = g-c(\theta-\theta_j)+\cdots. \qquad (3)$$

[1] A discussion of adiabatic methods in calorimetry is outside the scope of this paper.

[2] In the following account it is assumed that thermometric lags have been allowed for.

Quadratic and higher terms may be added if necessary; but with proper experimental precautions these terms usually are negligible.

Values of the "heating constant" h and the "heating coefficient" b obviously depend upon the nature of the particular experiment to which Eq. (2) is applied. The term $-c(\theta-\theta_j)$ in Eq. (3) is the analytical expression of Newton's law of cooling in which c is the so-called "cooling coefficient" of the calorimeter. Heat exchange may take place by evaporation, radiation, conduction, and convection. Evaporation is made negligible and the other losses are made small by proper construction of the calorimeter.[3] Under these circumstances Newton's law is appropriate for the small temperature ranges encountered in precision measurements. The discussion below will be restricted to cases where c and θ_j are constant throughout an experiment. The term g in Eq. (3) depends, for the most part, upon the rate of development of heat due to stirring. Since this rate is supposed constant, it is convenient to define an effective jacket temperature θ_k so that

$$f(\theta-\theta_j)=-c(\theta-\theta_k). \tag{4}$$

In calorimetric experiments, the source may supply heat to the calorimeter for a *limited* period of time as in cases of electrical heating or else for an *unlimited* period of time as in measurements of specific heats of solids. These two types of experiments in general require different methods for correcting the temperature rise of the calorimeter fluid.

In the first type of experiment, temperature readings are made at equal time intervals before the addition of heat (foreperiod), during the addition of heat (heating period), and afterwards (afterperiod). The change in temperature during these periods should not display discontinuities or marked unevenness. To illustrate these remarks, Fig. 1 is drawn for a case of electrical heating. Although numerous methods for making the correction in such experiments have been suggested, a perusal of Part I of this article will

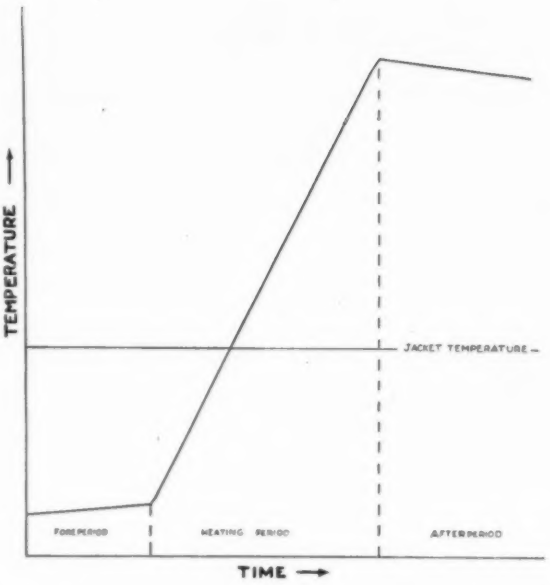


FIG. 1. A typical temperature-time graph of calorimetric data with a limited heating period.

reveal that all of these are modifications of a single general scheme.

The fundamental problem discussed in Part II is that in which heat energy is being liberated continually to the calorimeter fluid over an unlimited heating period. Data for such an experiment consist of temperature readings at equal time intervals before the unknown quantity of heat is supplied to the calorimeter (foreperiod) and during the time that the source supplies heat (heating period). The heating period has been considered to end when the temperature-time curve begins to obey, within experimental error, a pure exponential relation. Our analysis of this problem indicates that this assumption is not valid. Even though the amount of heat transferred from source to fluid per unit time continually diminishes, at the same time the rate of loss of heat to the surroundings diminishes as well; so that the first effect seldom becomes negligible with respect to the second. It is for this reason that the exponential character of the heating period after sufficient length of time does not imply pure Newtonian cooling. The straight line portion of the θ *vs.* t curve beyond the maximum as illustrated in Figs. 2 and 3 is not a cooling curve. For precision work, therefore, the method of Part I should not

[3] W. P. White, *The Modern Calorimeter* (The Chemical Catalogue Co., New York, 1928). H. C. Dickinson, Sci. Pap. Bur. Stand. **11**, 189 (1914). Dickinson ascribes one-fifth of the heat exchange to radiation and four-fifths to conduction and convection (p. 195). Actually convection is made small at the expense of increased conduction but such that the total correction term is as small as possible.

be applied to specific heat and bomb calorimetric data. Rowland originally recognized this difficulty and suggested a method based upon some knowledge of the heat function $F(\Theta-\theta)$. An examination of this and similar methods reveals fundamental errors. Consequently, in Part II we develop a method which we believe to be basically better.

Equation (2) offers a classification of these correction methods. In Part I experiments are considered for which $F(\Theta-\theta)$ is zero during foreperiod *and afterperiod* but is not otherwise specified. Part II deals with two cases in which $F(\Theta-\theta)$ is specified; in the first, h is constant and b is zero; in the second, h is zero and b is constant. Finally several graphical methods of making the correction are discussed critically in Part III.

Because a large number of symbols is necessary for the discussion of these correction methods, a table has been appended to facilitate the reading of the paper. The complete meaning of a given symbol must be obtained from the text, since the definition in the table is necessarily limited.

TABLE OF SYMBOLS

t', t, t''	time during foreperiod, heating period, and afterperiod
t_0, t_m	time at beginning, and at end of heating period
$t_{\max}$	time when calorimeter temperature is a maximum
$t_{\inf}$	time when calorimeter temperature-time curve has a point of inflection
$\theta', \theta, \theta''$	calorimeter temperature in foreperiod, heating period, and afterperiod
$\theta_a', \theta_a, \theta_a''$	"average" calorimeter temperature in each period
θ_0, θ_m	calorimeter temperature at beginning and at end of heating period
θ_j	jacket temperature
θ_k	effective temperature of surroundings (jacket temperature plus a correction term due to rate of stirring)
$\theta_{\max}$	maximum calorimeter temperature
$\theta_{\inf}$	temperature at point of inflection of calorimeter temperature-time curve
θ^*	ideal calorimeter temperature that would exist if there were no heat exchange with surroundings
Θ	temperature of source of heat
Θ^*	ideal temperature of source in absence of exchange with surroundings
$\Theta_0=\Theta_0^*$	temperature of source at beginning of heating period
v', v''	slope of foreperiod, and afterperiod curve
α	slope of calorimeter temperature-time curve at the point of inflection
a	cooling coefficient for the source of heat
b	heating coefficient for calorimeter
c	cooling coefficient of calorimeter
g	rate of increase of calorimeter temperature due to stirring
h	heating constant of calorimeter
p	abbreviation for $\frac{1}{2}(a+b+c)$
ϵ	abbreviation for $p-(p^2-ac)^{\frac{1}{2}}$

PART I

Limited heating period with heating function unspecified

Measurements of latent heats of fusion or vaporization or of the electrical equivalent of heat are examples of the class of experiments in which heat is supplied to the calorimeter during a limited time. If the time rate of change of calorimeter temperature cannot be formulated explicitly for such experiments, then the methods discussed in this section are used.

During the foreperiod of such an experiment, $l+1$ successive temperature readings θ_0', θ_1', $\cdots$, θ_l' are made at the times t_0', t_1', $\cdots$, t_l'. At the time $t_l'=t_0$, the quantity of heat to be measured is allowed to begin to flow into the calorimeter. During the ensuing heating period, $m-1$ readings of temperature θ_1, θ_2, $\cdots$, θ_{m-1} are obtained at times t_1, t_2, $\cdots$, t_{m-1}. Finally, at time $t_m=t_0''$ the addition of heat is stopped and during an afterperiod $n+1$ readings of temperature θ_0'', θ_1'', $\cdots$, θ_n'' are taken at times t_0'', t_1'', $\cdots$, t_n''. Usually the time intervals between successive readings are kept equal in length throughout the whole experiment.

Clearly the observed rise in temperature of the calorimeter is just the difference in temperatures at the beginning and end of the heating period, namely $\theta_m-\theta_0$. During this rise, however, heat exchanges with the surroundings occur so that a correction term C must be added to get the change in temperature $\Delta\theta$ of the calorimeter due to the unknown quantity of heat alone. Symbolically,

$$\Delta\theta=\theta_m-\theta_0+C. \tag{5}$$

We wish to examine methods of evaluating C from the experimental data. Of course, the accuracy of the value for $\Delta\theta$ depends not only

upon how well C can be estimated, but also upon the precision with which θ_m and θ_0 are found from the data.

Since the heating function need not be specified explicitly let $F(\Theta-\theta)=F(t)$ so that Eq. (1) becomes

$$d\theta/dt=F(t)-c(\theta-\theta_k). \tag{6}$$

None of the heat being measured is allowed to enter the calorimeter during foreperiod and afterperiod; hence during these periods $F(t)=0$ and, upon integrating Eq. (6),

$$\theta'=\theta_k+(\theta_i'-\theta_k)\exp[-c(t'-t_i')] \tag{7}$$

and

$$\theta''=\theta_k+(\theta_0''-\theta_k)\exp[-c(t''-t_0'')].$$

For the heating period, on the other hand, $F(t)\neq 0$ so that

$$\theta=\theta_0+\int_{t_0}^{t}F(t)dt-\int_{t_0}^{t}c(\theta-\theta_k)dt. \tag{8}$$

For $t=t_m$ the first integral on the right of Eq. (8) is seen to be just $\Delta\theta$ and the second integral is the correction term

$$C=\int_{t_0}^{t_m}c(\theta-\theta_k)dt. \tag{9}$$

An average temperature θ_a for the heating period is now defined such that

$$\theta_a=(t_m-t_0)^{-1}\int_{t_0}^{t_m}\theta dt. \tag{10}$$

Suppose the times to be reckoned in units equal to the time interval used; then $t_m-t_0=m$ and

$$C=c(\theta_a-\theta_k)m. \tag{11}$$

Difficulties in finding c, θ_a, and θ_k from the observations arise mainly because the temperature readings are made at finite time intervals.

It should be possible to get c and θ_k from either foreperiod or afterperiod data when analyzed by Eqs. (7). To get both c and θ_k from the data of one period requires, however, sufficient accuracy to detect variations of $d\theta/dt$ with temperature within this period. Usually these variations are so small that the graphs representing foreperiod and afterperiod data may well be fitted with straight lines. In such cases both sets of data are used to evaluate c and θ_k.

Evidently the slopes v' and v'' of these straight lines may be considered to equal the slopes of the exponentials at temperatures θ_a' and θ_a'' within foreperiod and afterperiod, respectively. Thus

$$v'=(d\theta'/dt')_a=-c(\theta_a'-\theta_k) \tag{12}$$

and

$$v''=(d\theta''/dt'')_a=-c(\theta_a''-\theta_k),$$

where θ_a' and θ_a'' are appropriate "average" temperatures.

If v', v'', θ_a', and θ_a'' can be obtained from the experimental data then the constants c and θ_k may be computed from Eqs. (12); so that

$$c=(v'-v'')/(\theta_a''-\theta_a') \tag{13}$$

and

$$\theta_k=(v'\theta_a''-v''\theta_a')/(v'-v''). \tag{14}$$

Finally these values of c and θ_k are substituted in Eq. (11) to get

$$\begin{aligned}C&=-mv'-(v'-v'')(m\theta_a'-m\theta_a)/(\theta_a''-\theta_a') \\ &=-mv''-(v'-v'')(m\theta_a''-m\theta_a)/(\theta_a''-\theta_a').\end{aligned} \tag{15}$$

Many of the methods proposed in the literature agree in Eqs. (15) but differ in the manner of computing v', v'', θ_a', θ_a, θ_a'', θ_0, and θ_m.

In view of the assumed linearity of the foreperiod data it is clear that by fitting a straight line to these data both v' and θ_a' may be determined. Application of the method of least squares yields the result that

$$v'=6\{l(l+1)(l+2)\}^{-1}\sum_{r=0}^{r=l}(2r-l)\theta_r' \tag{16}$$

and

$$\theta_a'=(l+1)^{-1}\sum_{r=0}^{r=l}\theta_r', \tag{17}$$

where θ_r' represents the observed temperature at time t_r' in the foreperiod. Similar relations may be found for v'' and θ_a''. Equations equivalent to these are used by Roth.[4]

A second method that has been used to obtain v' and θ_a' may be understood by supposing the end points of the intervals within the foreperiod to be joined by straight lines. It is now assumed

[4] W. A. Roth, *Thermochemie* (Sammlung-Göschen, Berlin and Leipzig, 1932), p. 26. Note that Roth applies his method to bomb calorimetric experiments for which the heating period is not limited.

that v' is equal to the arithmetic mean of the slopes of these lines and that θ_a' is equal to the arithmetic mean of the temperatures at their midpoints. Evidently this method may not yield an adequate value for v' since the value obtained depends entirely on the observed temperatures at the ends of the foreperiod. Of course, v'' and θ_a'' may be determined in a similar manner. Schultes and Nübel[5] propose this method and claim that the value of θ_a' so determined is superior to that given by Eq. (17). Regnault and Pfaundler[6] find v' in the above manner and use Eq. (17) for θ_a'.

Once suitable values for v', v'', θ_a', and θ_a'' have been found, then θ_0 and θ_m may be computed from the relations

$$\theta_0=\theta_a'+\tfrac{1}{2}lv', \tag{18}$$

$$\theta_m=\theta_a''-\tfrac{1}{2}nv'' \tag{19}$$

as suggested by Roth and by Schultes-Nübel.

Three methods have been proposed for computing θ_a. First, the mean temperature in the heating period is satisfactory if the data in this period obey a linear relationship. Thus

$$\theta_a=(m+1)^{-1}\sum_{r=0}^{r=m}\theta_r. \tag{20}$$

Second, Regnault-Pfaundler, Roth, and Schultes-Nübel all assume that θ_a is the average value of the temperatures at the midpoints of the straight lines between successive observed points in the heating period. This averaging leads to a slightly different expression, namely

$$\theta_a=m^{-1}\{\sum_{r=1}^{r=m-1}\theta_r+\tfrac{1}{2}(\theta_0+\theta_m)\}. \tag{21}$$

Finally, White[7] suggests that Simpson's rule may be used to get θ_a. For instance, if m is even

$$\theta_a=(3m)^{-1}\{(\theta_0+\theta_m)+4(\theta_1+\theta_3+\cdots)+2(\theta_2+\theta_4+\cdots)\}. \tag{22}$$

It must be remembered that θ_a is defined by Eq. (10) in terms of the area under the θ *vs.* t curve for the heating period. For an approximately linear curve, the simple expressions in Eqs. (20) and (21) are quite adequate; for some experiments more laborious ways of estimating the area may be needed.

It is obvious that simplifying approximations in the calculation of any one of the seven constants can often be made for a given set of experiments. Although any such simplifications can be adopted only for particular problems, some suggested approximations are of sufficiently general interest to deserve mention here.

Schultes and Nübel[5] observe that if v' and v'' are taken to correspond to θ_0 and θ_m instead of θ_a' and θ_a'' and if θ_a is given by Eq. (21) then

$$C=-mv'-(v''-v')\{\tfrac{1}{2}+(\theta_m-\theta_0)^{-1}\sum_{r=1}^{r=m-1}(\theta_r-\theta_0)\}. \tag{23}$$

Langbein[8] notes that if the rise in temperature in the heating period takes place so rapidly that the elevation occurs during the first interval[9] and if $\theta_1=\theta_2=\cdots=\theta_m$, $\theta_a'=\theta_0$, and $\theta_a''=\theta_m$ then

$$C=-mv''-\tfrac{1}{2}(v'-v''). \tag{24}$$

On the other hand, Moser[10] observes that Eq. (15) may be written

$$C=-mv''-G(v'-v''), \tag{25}$$

where G is nearly constant for a particular apparatus and a given class of experiments. Actual values for G (from *bomb calorimetric* experiments) varied from 1.0 to 1.5.

White[7] points out that if $\theta_a''-\theta_a$ is small it is convenient to write Eq. (15)

$$C=-mv'-mk(\theta_a''-\theta_a). \tag{26}$$

Then, if k may be considered constant for a set of experiments, a single preliminary determination of k eliminates the necessity of taking foreperiod data for each run.

Graphical analogs of some of the foregoing analytical methods are discussed in Part III.

Part II

Heating function specified

Sometimes the unknown quantity of heat is supplied in such a manner that the time rate of

[5] W. Schultes and R. Nübel, Die Wärme **58**, 15 (1935); Brennstoff-Chemie **15**, 466 (1934).

[6] L. Pfaundler, Ann. d. Physik **205**, 102 (1866). Griffiths' arithmetical method is equivalent to this one. See Glazebrook's *Dictionary of Applied Physics*, Vol. 1, p. 61.

[7] W. P. White, J. Am. Chem. Soc. **48**, 1146 (1926).

[8] F. Stohmann, C. Kleber and H. Langbein, J. prakt. Chemie **147**, 503 (1889).

[9] Strictly this is impossible when the heating period is unlimited. Langbein applies his method to the class of experiments discussed in Part II.

[10] H. Moser, Physik. Zeits. **37**, 529 (1936).

change of calorimeter temperature can be approximated by a simple analytical function. This fact allows another approach to the problem of calculating the temperature correction.

To illustrate this approach, we shall consider the case in which the heating function $F(\Theta-\theta)$ is constant over a limited time (the heating period) and zero otherwise. This situation may occur, for example, in measuring the electrical equivalent of heat. Eq. (1) then becomes

$$d\theta/dt=h-c(\theta-\theta_k). \tag{27}$$

Its solution may be written

$$\theta-\theta_k=h/c-(h/c-\theta_a+\theta_k)\exp[-c(t-t_a)], \tag{28}$$

where the point (t_a, θ_a) lies in the heating period. The constants c and θ_k can be evaluated from foreperiod and afterperiod data as described in Part I.

The expansion of the exponential in Eq. (28) about a properly chosen value of ct_a through the linear term is sufficiently accurate for many experiments. Expand, insert the end values of the heating period, and eliminate t_a between the resulting two equations to get

$$h(t_m-t_0)=\theta_m-\theta_0+c(\theta_a-\theta_k)(t_m-t_0). \tag{29}$$

The temperature θ^* which the calorimeter and contents would have had in the absence of stirring and of heat exchange with the surroundings is the solution of Eq. (27) with the correction term omitted, namely

$$\theta^*-\theta_0^*=h(t-t_0), \tag{30}$$

where $\theta^*=\theta_0^*=\theta_0$ when $t=t_0$. Finally, the desired corrected temperature rise $(\theta_m^*-\theta_0^*)$ is obtained from Eqs. (29) and (30). Thus

$$\theta_m^*-\theta_0^*=\theta_m-\theta_0+c(\theta_a-\theta_k)(t_m-t_0). \tag{31}$$

Equation (31) has the same form as the relation obtained from Eqs. (5) and (11) in Part I. Generally the term h in Eq. (27) predominates over the small correction term so that the θ *vs.* t curve is nearly linear over the heating period. It is then sufficiently accurate to let $\theta_a=\frac{1}{2}(\theta_0+\theta_m)$ so that

$$\theta_m^*-\theta_0^*=\theta_m-\theta_0+\tfrac{1}{2}c(t_m-t_0)(\theta_0+\theta_m-2\theta_k). \tag{32}$$

Jaeger and von Steinwehr[11] and Hoare[12] obtain this relation. Of course, better approximations may be found.

In his famous investigation on the mechanical equivalent of heat, Rowland[13] develops a method in which the nature of the heat exchange between a hot body immersed in the calorimeter fluid and the fluid is assumed to be known. Hoare[12] gives a more mathematical treatment of this problem. The Rowland and the Hoare treatments, however, are based on a fallacious assumption and lead to erroneous results as will be shown later.[14] Because of this fact and because the Regnault-Pfaundler treatment is improper for experiments on specific heats, bomb calorimetry, and the like, we believe an extended discussion of the Rowland problem is justified.

If results good to one-tenth of one percent or better are desired, then great care must be exercised in determining the corrected temperature rise. In such cases not only are the Regnault-Pfaundler and Rowland-Hoare treatments inadequate but also Dickinson's method,[15] which is recommended by the ASTM standards, becomes seriously questionable. Of course, if errors of one-half of one percent or worse are acceptable then almost any of these methods may be applied.

We make three fundamental assumptions. *First*, at the initial time t_0 the charge immersed in the calorimeter is at a uniform elevated temperature and the calorimeter fluid is at a uniform lower temperature.[16] As time goes on, the charge temperature decreases and the calorimeter temperature increases, but always in such a way that both temperatures are uniform over the entire bodies in question. In practice one may conceive of a thin conducting shell enclosing the charge-fluid boundary and having a heat capacity small compared with that of the charge and that of the fluid. Across this shell there is a temperature gradient proportional to the temperature

[11] W. Jaeger and H. von Steinwehr, Ann. d. Physik **64**, 305 (1921), see p. 334.

[12] F. E. Hoare, Phil. Mag. **29**, 52 (1940), see p. 57.

[13] H. A. Rowland, *Physical Papers* (The Johns Hopkins Press, Baltimore, 1902), p. 400. The problem presented here will be termed the Rowland problem.

[14] See page 565.

[15] H. C. Dickinson, reference 3, pp. 230–234. Also see p. 567 of this article.

[16] Of course, the charge temperature may be lower than that of the calorimeter. This situation does not often occur.

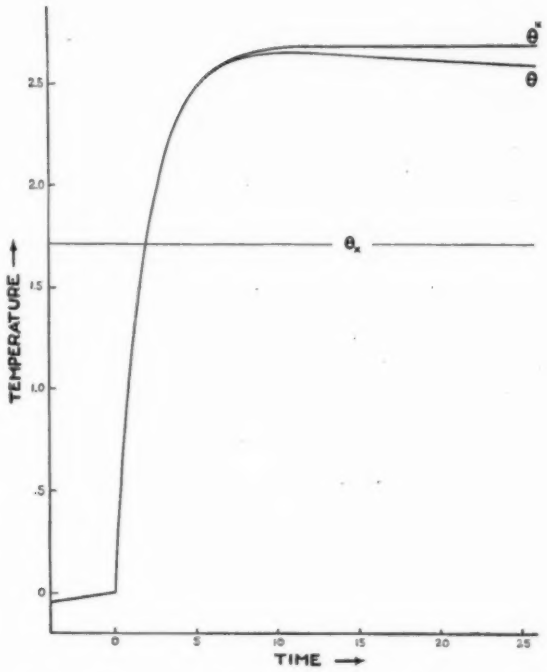


FIG. 2. Plot of the data for an unlimited heating period (see text). Note that an afterperiod does not exist for such data. In this figure θ_x should read θ_k.

difference between charge and calorimeter fluid. Dickinson[17] and White[18] imply that this assumption is not wholly incompatible with the true situation. More likely models are readily visualized but these require too complicated analyses. At any rate, other modifications are introduced due to the act of immersing the charge or the process of igniting the bomb.

Second, the charge temperature Θ is considered to change at a time rate which is proportional to the temperature difference between charge and calorimeter; so that

$$d\Theta/dt = -a(\Theta-\theta), \tag{33}$$

in which a, the proportionality constant, is a "cooling coefficient" for the charge. Actually the rate is undoubtedly a more complicated function of $(\Theta-\theta)$; the right-hand member of Eq. (33) may be taken as the second term in a power series expansion in which the first term and terms higher than the second are negligible. Note that $d\Theta/dt$ does not depend explicitly upon θ_k.

Third, the time rate of change of the calorimeter temperature is supposed to be given by

$$d\theta/dt = b(\Theta-\theta) - c(\theta-\theta_k). \tag{34}$$

Here we have expanded the function $F(\Theta-\theta)$ into a power series and neglected the first and the higher terms. That the first terms in the two series contained in Eqs. (33) and (34) are zero seems plausible when we remember that after a long time both $d\Theta/dt$ and $d\theta/dt$ should approach zero. Whether the terms higher than the second are always negligible may be seriously questioned. But in view of the "idealness" of the first assumption, we may accept the second and third assumptions as adequate.

Evidently, if no extraneous heat exchanges occur, then from the above assumptions with the correction term omitted, the temperature of the calorimeter will rise exponentially as shown by the upper curve (the θ^* curve) in Fig. 2. On the other hand, if exchanges do occur as in the case of observed data, then (for $\theta_0<\theta_k<\theta_{\max}$) the temperature θ rises at first more rapidly than θ^* and later more slowly than θ^*; it reaches a maximum value $\theta_{\max}$ and finally falls off *slowly* toward θ_k. This is shown by the lower curve (the θ curve) in Fig. 2. If such observed data be fitted by the analytical expression of a θ *vs.* t curve, constants of the corresponding θ^* curve are thereby evaluated so that the desired corrected temperature rise $(\theta_x^*-\theta_0^*)$ may be calculated.

Equations (33) and (34) yield the standard differential equation

$$d^2\theta/dt^2 + 2p\,d\theta/dt + q\theta = q\theta_k, \tag{35}$$

where $2p\equiv a+b+c$ and $q\equiv ac$. The solution may be written

$$\theta = \theta_k + Ae^{-\epsilon t} + Be^{-(2p-\epsilon)t}, \tag{36}$$

where $\epsilon\equiv p-(p^2-q)^{\frac{1}{2}}$, $p^2>q$,[19] and A and B are constants of integration.

Clearly the straight line portion of the θ-curve beyond the maximum is not a cooling curve. Since $2p>\epsilon$, after a sufficient length of time, the third term in Eq. (36) is negligible with respect to the other terms and the θ *vs.* t curve then depends upon $e^{-\epsilon t}$ rather than upon e^{-ct}. Figure 3 shows

[17] H. C. Dickinson, reference 3, p. 203.
[18] W. P. White, reference 3, p. 88.

[19] Since a, b, and c are positive, solutions of Eq. (35) in which $p^2\leq q$ are inadmissible.

how much the slope in this region may differ from the slope of a cooling curve. From Eqs. (33) and (34) it follows that

$$v'' = (d\theta/dt)_{\text{cooling curve}}$$

$$= (2p/a)(d\theta/dt)_{\theta\ \text{curve}} = (a+b+c)\alpha/a \quad (37)$$

at the point of inflection which is usually in the straight line portion of the actual curve. Evidently the two slopes are practically equal only if $(b+c) \ll a$ which is rarely satisfied.[20] Hence it is improper to use data after the maximum temperature has been reached as afterperiod data from which to obtain c and θ_k in the manner discussed in Part I.

To obtain the corrected temperature rise, the equation resulting from Eqs. (33) and (34) with $c=0$ is solved. Thus

$$\theta^* = A^* + B^* e^{-(a+b)t}, \quad (38)$$

where A^* and B^* are constants of integration. If, at $t=0$, $\theta_0 = \theta_0^*$ and $\Theta_0 = \Theta_0^*$, then Eq. (34) with $c=0$ and Eq. (38) yield[20]

$$\theta_x^* - \theta_0^* = b(\Theta_0 - \theta_0)/(a+b). \quad (39)$$

In general it is difficult to use Eq. (39) in this form. However, it is easy to develop an expression for the corrected temperature rise in terms of $(\theta_0 - \theta_k)$ and $(\theta_{\max} - \theta_k)$.[21] Using Eqs. (34), (39), and the expressions resulting from Eq. (36) evaluated at $t=0$ and again at $t=t_{\max}$:

$$\theta_x^* - \theta_0^* = \{\epsilon(a+b)^{-1} - 1\}(\theta_0 - \theta_k)$$

$$+ (2p-\epsilon)(a+b)^{-1}(\theta_{\max} - \theta_k)\exp(\epsilon t_{\max}). \quad (40)$$

Additional useful relations are

$$xe^{-x} = ye^{-y} = -\alpha(t_{\max} - t_k)/(\theta_{\max} - \theta_k) \quad (41)$$

and

$$t_{\inf} - t_k = 2(t_{\max} - t_k), \quad (42)$$

where $x \equiv \epsilon(t_{\max} - t_k)$, $y \equiv (2p-\epsilon)(t_{\max} - t_k)$, α is the slope of the θ curve at the time $t_{\inf}$ where it has its point of inflection, and t_k is the time at which $\theta = \theta_k$.

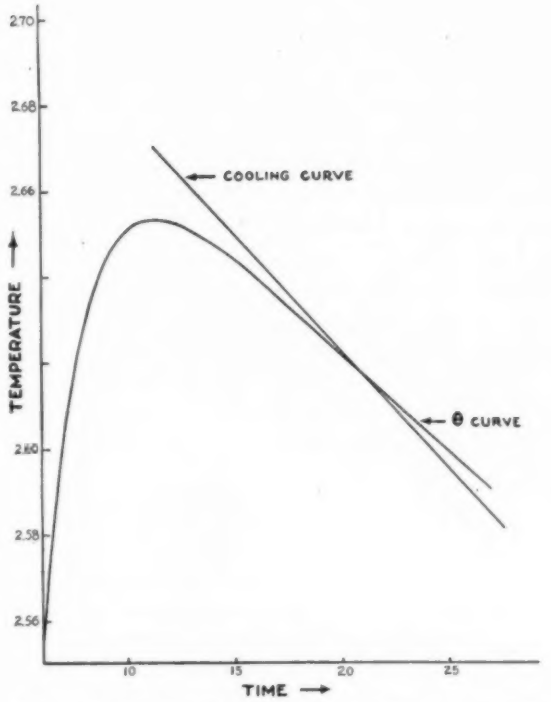


FIG. 3. Upper part of the θ curve in Fig. 2. Note that the straight line portion beyond the maximum is not a cooling curve. A cooling curve which is appropriate to the given data ($c = 0.006$ int^{-1}) has been drawn through the point of inflection.

For many sets of data wherein values of θ have been observed over all three periods, the quantities θ_0, $\theta_{\max}$, $t_{\max}$, α and v' may be obtained directly. Although readings of θ and t in the neighborhood of the point of inflection may be fairly precise, the point itself is difficult to locate since the curve is so nearly linear in this region. Values of θ over the principal rise in the heating period are often not precise and may be affected by lags in making the readings as well as other factors mentioned above. Hence application of the theory to such data is difficult and sometimes ambiguous. By trial and error, values of θ_k and t_k are found which satisfy not only Eq. (41) but also Eq. (36) over a large portion of the heating period and the "afterperiod." If an unambiguous value of θ_k can be so determined; then c may be found from foreperiod data by Eq. (12) of Part I, ϵ and $2p$ from Eqs. (41), and finally the corrected temperature rise from Eq. (40).

[20] Apparently a and b are inversely proportional to the heat capacities of charge and calorimeter, respectively. From the well-known equation of mixtures there then results

$$a/b = (\Theta_0 - \Theta_\infty^*)/(\theta_\infty^* - \theta_0),$$

where $\Theta_\infty^* = \theta_\infty^*$ is the final temperature of the mixture. Equation (39) follows at once. Evidently a/b usually is *not very large.*

[21] An equation involving α instead of $(\theta_{\max} - \theta_k)$ is readily obtained. Other forms of possible use in certain cases may be suggested by examination of the various relations given in the text.

In another type of experiment[22] the quantities $\theta_0=\theta_k$, t_0, α, and sometimes rough values of θ_{max} and t_{max} are found. Readings of sufficient number are taken so that the exponential character of the "afterperiod" is clearly evident. On the other hand, no data are obtained over the principal rise of the heating period. In this case, values of $(\theta-\theta_0)$ over the "afterperiod" are fitted with a pure exponential since the third term on the right-hand side of Eq. (36) is negligible in this region. Thus ϵ and A are determined. By trial and error values of θ_{max} and t_{max} are found so that both Eqs. (36) and (41) are satisfied and thus $(2p-\epsilon)$ is obtained. If c is found from foreperiod data, then Eq. (41) gives the corrected temperature rise.

A better fit for data of the first type can be made if c has been determined for a given apparatus from preliminary experiments. Then θ_k can be found from foreperiod data. If readings during the heating period are sufficiently accurate, t_k may now be read directly from the θ curve. Otherwise a plot of $\ln(\theta-\theta_k)$ for the "afterperiod" is used to determine ϵ; in this case, Eq. (41) then gives t_k. On the other hand, if c *is small* and the ratio of the heat capacities of charge and calorimeter can be estimated, then Eq. (37) gives the approximate slope of the cooling curve in the neighborhood of the inflection point.[23] Both c and θ_k can now be computed by the method of Part I.

Even if c and θ_k are known, it is not, in general, correct to use the method of Part I to compute the desired θ^* curve from the observed θ curve. Consideration of Eqs. (33) and (34) shows that

$$\theta-\theta^*=-\int_0^t c(\theta-\theta_k)dt+(b/a)(\Theta^*-\Theta). \quad (43)$$

Since Θ^* and Θ are not equal over the entire heating period, it is incorrect to neglect the second term on the right-hand side of Eq. (43).

It is convenient to mention at this point the approximations in the treatments of Rowland and Hoare. The latter deduces an expression for θ_x^* by neglecting in Eq. (43) the term involving Θ^* and Θ and by replacing θ under the integral by θ^*.[24] Since, for small values of c, the effect of the second approximation is not important, Hoare's equation usually gives too high a value for θ_x^*. However, instead of using the equation directly (which is difficult in practice), Hoare develops a graphical solution similar to that originally proposed by Rowland. This graphical solution involves additional approximations which sometimes override the approximations in the derivation of the equation, so that the net result is too low a value for θ_x^*.

For many experiments where $(c-\epsilon)$ is very small compared with $(a+b)$ the method based upon Eq. (36) may be simplified. The initial temperature θ_0 is supposed to be equal to θ_k; that is, with the stirrer in operation, the calorimeter and fluid are allowed to come to constant temperature before the charge is dropped in or before the bomb is ignited. The first term in Eq. (40) is then equal to zero.

Over the "afterperiod" the last term in Eq. (36) is negligible to the approximation we are considering; hence the temperatures at the inflection point (θ_{inf}) and at $t_{max}(\theta_x)$ will be given by

$$\theta_{inf}-\theta_0=A\exp(-\epsilon t_{inf}) \quad (44)$$

and

$$\theta_x-\theta_0=A\exp(-\epsilon t_{max}), \quad (45)$$

respectively. Remembering Eq. (42) there results

$$\exp(\epsilon t_{max})=(\theta_x-\theta_0)/(\theta_{inf}-\theta_0). \quad (46)$$

Now if $(c-\epsilon)\ll(a+b)$, then Eqs. (40) and (46) yield

$$\theta_\infty^*-\theta_0^*=(\theta_x-\theta_0)(\theta_{max}-\theta_0)/(\theta_{inf}-\theta_0). \quad (47)$$

If the "afterperiod" is linear, Eq. (47) becomes

$$\theta_x^*-\theta_0^*=(\theta_{max}-\theta_0)\times\{1-(\alpha t_{max})/(\theta_{inf}-\theta_0)\}. \quad (48)$$

A simple graphical construction for this method is given in Part III.

It is, of course, difficult to give an estimate of the amount of error involved in any proposed method of evaluating the corrected temperature change, since this error depends upon the con-

[22] F. M. Jaeger and E. Rosenbohm, Proc. K. Akad. Amst. **30.8**, 965 (1927).

[23] An examination of Eqs. (37) and (41) shows that an error in α has a greater influence in the first equation. A correction of the observed average "afterperiod" slope can be found from the relation

$$\alpha=\alpha_{Av}-\tfrac{1}{2}\{(d^2\theta/dt^2)_{inf}(t-t_{inf})^2{}_{Av}\}.$$

[24] The first of these approximations is equivalent to replacing the assumption expressed in Eq. (33) by

$$d\Theta/dt=-a(\Theta-\theta)-c(\theta-\theta_k).$$

stants of the particular apparatus used. In our search through the literature we failed to discover published data of sufficient extent to allow an unambiguous determination of the constants in Eq. (36) and so to afford a numerical comparison of the result of this equation with the results from other proposed methods. Consequently, we shall illustrate the errors involved by a set of artificial data computed by assuming reasonable values of the constants in Eq. (36).

The curves in Figs. 2 and 3 are plots of Eqs. (36) and (38) with the chosen values $\theta_0=0°$, $t_k=2$ time intervals, $\epsilon=0.005$ int^{-1}, $2p=0.505$ int^{-1}, and $c=0.006$ int^{-1}. Other quantities in the above discussion then have the values: $a=0.4167$ int^{-1}, $b=0.0823$ int^{-1}, $\alpha=-0.004510$ deg./int, $\theta_k=1.7082°$, $\theta_{\max}=2.6532°$, $\theta_{\inf}=2.6193°$, $t_{\max}=11.3034$ int, and $t_{\inf}=20.6068$ int. For this data the corrected temperature rise is $\theta_x{}^*=2.6931°$.

The "corrected infinity temperature" θ_I calculated by the method of Part I (assuming the "afterperiod" data to give a true cooling curve) should be less than $\theta_x{}^*$, since in this calculation the effect of the transfer of heat from the charge to the calorimeter fluid during the "afterperiod" is neglected. The corresponding temperature θ_H from Hoare's equation should be too high; on the other hand, the value θ_{R-H} computed by the graphical method of Rowland and of Hoare may well be too low. A similar analysis shows that the value θ_D calculated by Dickinson's method may also be expected to differ from $\theta_x{}^*$. The results of such errors in the case of our artificial data are indicated in Table I. The approximate method suggested on page 565 should not be used for the above data. Application of this method to similar data (for which $\theta_0=\theta_k$) gave a result only 0.1 percent too high.

In conclusion, it must be noted again that the differences among the results of various correction methods are significant only in precision work where a final error of not more than 0.1 percent is desired.

TABLE I. *Results of applying various correction methods to illustrative data.*

	"Corrected Infinity Temperatures"	% Error
$\theta_\infty{}^*$	2.6931	0
θ_I	2.6859	0.27 low
θ_H	2.7000	0.26 high
θ_{R-H}	2.6884	0.17 low
θ_D	2.6979	0.18 high

PART III

Graphical methods

Graphical methods for determining the corrected rise in temperature in calorimetric experiments are designed to avoid undue and tedious calculations. To question the relative accuracy of the algebraic and graphical methods in terms of the size of the graph paper needed appears trivial to us. Usually, the *total* corrected temperature rise is determined directly by graphical methods so that large errors are likely to creep in. Occasionally, only the correction term is found by graphical methods; in such cases, the evaluation of the uncorrected temperature rise remains an important problem.

In the original Rumford[25] method, the value of the initial temperature θ_0 is chosen on the basis of preliminary experiments so that $\theta_k=\frac{1}{2}(\theta_0+\theta_{\max})$. This method has been modified in such a way that θ_0 is chosen so as to make the areas A_1 and A_2 in Fig. 4 equal. The corrected temperature is now supposed to be $(\theta_{\max}-\theta_0)$. Of course, it is not always true that an observed $\theta_{\max}$ is the final temperature of the heating period; then it is difficult to apply this method. Furthermore, the time required to find an appropriate value of θ_0 may be greater than that to calculate a correction by methods suitable for use with an arbitrary value of θ_0.

Regnault and Pfaundler[6] suggest a semi-graphical method[26] based upon their analytical treatment (see Part I). A straight line is drawn through the points (v', θ_a') and (v'', θ_a'') on graph paper with slope-temperature coordinates. Values of v are read off this curve at temperatures which are equal to the arithmetic means for the time intervals during the heating period. The sum of these values gives the desired correction. Evidently the criticisms given in Part I apply here.

Jaeger and von Steinwehr[27] determine the constants c and θ_k from the fore- and afterperiod

[25] *The Complete Works of Count Rumford* (Estes and Lauriant, Boston, 1870–1875), Vol. II, p. 376.

[26] For a similar construction see P. E. Klopsteg, J. Opt. Soc. Am. **13**, 589 (1926).

[27] W. Jaeger and H. von Steinwehr, Zeits. f. physik. Chemie **53**, 153 (1905).

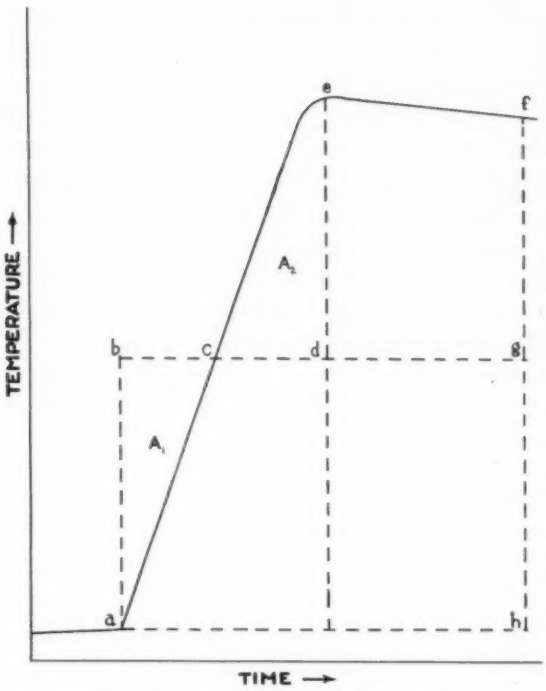


FIG. 4. Illustration of the Rumford and Jaeger-von Steinwehr graphical methods.

data as described in Part I. Areas *abc* and *cefgd* (Fig. 4) are measured with a planimeter or by counting squares. The correction term C is obtained by multiplying the difference between these areas by c. If the rise *ac* is very rapid, so that *abc* is small, they suggest that area *acdgh* may be measured and subtracted from the rectangular area *abgh*. It is difficult for us to see that a more accurate evaluation of the required area is gotten in this manner. They also show that the position of point *f* is immaterial as long as it is located on the *linear* portion of the afterperiod curve.

In many methods a time t_x is assumed to exist such that the difference in readings on the extrapolated foreperiod and afterperiod curves at t_x is equal to the desired corrected temperature rise. For that matter, once t_x has been found C may be calculated by means of the relation

$$C=-v'(t_x-t_0)-v''(t_m-t_x). \qquad (49)$$

We wish to emphasize that a tacit assumption is made in all these methods: namely, that the rate of heat exchange during the foreperiod *and* the afterperiod depends only upon the cooling constant c and the temperature difference $(\theta-\theta_k)$. These methods, therefore, are not to be applied to experiments in which the heating period is unlimited.

Ferry[28] proposes that t_x should be taken at θ_j on the heating period part of the curve; Eucken[29] suggests that θ_k should be used instead. Needless to say, in general, neither of these methods is correct. Roth[30] uses a value for t_x corresponding to the mean temperature $\frac{1}{2}(\theta_0+\theta_m)$. Smith[31] erects a perpendicular at such a time that the area between the extrapolated afterperiod curve, the latter part of the heating period curve, and this line is equal to the area between the extrapolated foreperiod curve, the earlier part of the heating period curve, and this same line. The methods of Roth and Smith are nearly equivalent to the Regnault-Pfaundler method[32] and ought not be used for experiments employing the method of mixtures.

In Dickinson's method[3] the time t_x is obtained in the following manner. On the plot of the calorimetric data lines are drawn at θ_0 and θ_m parallel to the time axis. Then a perpendicular to these lines is erected at such a place that the area between the θ_m line, the latter part of the heating curve, and the perpendicular is equal to the area between the θ_0 line, the earlier part of the heating curve, and the perpendicular. That this procedure yields a proper value for t_x may be readily shown.

For equality of the two areas evidently

$$\int_{t_0}^{t_x}(\theta_0-\theta)dt+\int_{t_x}^{t_m}(\theta_m-\theta)dt=0. \qquad (50)$$

But

$$\int_{t_0}^{t_x}(\theta_0-\theta)dt=\int_{t_0}^{t_x}(\theta_0-\theta_k)dt-\int_{t_0}^{t_x}(\theta-\theta_k)dt$$

and (51)

$$\int_{t_x}^{t_m}(\theta_m-\theta)dt=\int_{t_x}^{t_m}(\theta_m-\theta_k)dt-\int_{t_x}^{t_m}(\theta-\theta_k)dt.$$

[28] Glazebrook's *Dictionary of Applied Physics*, Vol. 1, p. 62.
[29] A. Eucken, "Energie und Wärmeinheit," *Handbuch der Experimentalphysik*, Vol. 8, p. 50.
[30] Reference 4, p. 30.
[31] E. S. Ferry, *A Handbook of Physics Measurements* (John Wiley and Sons, New York, 1929), Vol. 2, p. 77.
[32] Roth's method is identical with that of Regnault-Pfaundler if $\theta_a'=\theta_c$, $\theta_a''=\theta_m$, and $\theta_a=\frac{1}{2}(\theta_0+\theta_m)$. Smith's method agrees with the Regnault-Pfaundler formulation if θ_a' and θ_a'' are replaced by the temperatures at times $\frac{1}{2}(t_i+t_x)$ on the extrapolated foreperiod curve and $\frac{1}{2}(t_x+t_m)$ on the extrapolated afterperiod curve, respectively.

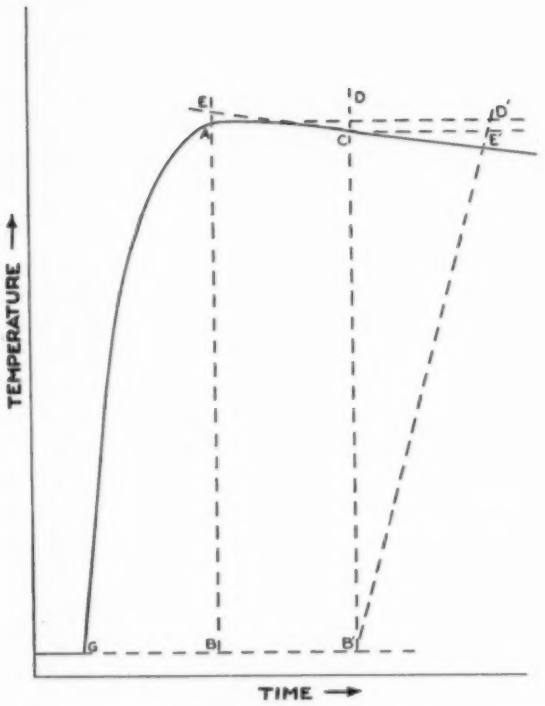


FIG. 5. A simple graphical analog of the analytical method described in Part II. It is to be used for bomb calorimetric and similar experiments.

Hence the correction term is just

$$C=-c\int_{t_0}^{t_m}(\theta-\theta_k)dt=-c(\theta_0-\theta_k)(t_x-t_0)$$
$$-c(\theta_m-\theta_k)(t_m-t_x). \quad (52)$$

The constants c and θ_k are found as in Part I.

Dickinson attempts to apply his method to bomb calorimetric experiments and derives an analytical expression for this purpose. Essentially he supposes that $t_0=0$ and that

$$\int_0^{t_x}(\theta^*-\theta_0^*)dt=\int_{t_x}^{\infty}(\theta_\infty^*-\theta^*)dt. \quad (53)$$

Insert Eq. (38) with A^* and B^* evaluated and integrate. There results that $t_x=(a+b)^{-1}$. Now solve Eq. (38) for θ_x^* to get $(\theta_x^*-\theta_0^*)=0.632\times(\theta_\infty^*-\theta_0^*)$. Dickinson replaces θ_∞^* with θ_m and considers the other temperatures to be on the θ-curve; the value of t_x corresponding to θ_x is then used in Eq. (53). For *his apparatus* Dickinson finds that 0.60 gives better results than does 0.632. It is this method which is recommended by the ASTM as a standard for finding the temperature correction[33] in bomb calorimetry.

None of the above methods is strictly applicable to specific heat or bomb calorimetric experiments of any kind. The Rowland-Hoare graphical scheme is designed to be used especially for specific heat experiments by the method of mixtures. The involved construction suggested by these men is based upon an analysis which we have already shown to be faulty (Part II); and, therefore, we propose a somewhat different method based upon our analysis of the Rowland problem.

Suppose the initial temperature of the calorimeter θ_0 just equals θ_k. On a temperature-time graph of the observed data (Fig. 5) draw lines parallel to the time axis through points G (at θ_0) and A (at θ_{max}). Drop a perpendicular from A to B on the line through G and with dividers mark off $BB'=GB$. Now erect a perpendicular to GB' at B' and let it intersect the θ *vs.* t curve in C. Draw a line through C parallel to the time axis. Extrapolate the "afterperiod" part of the θ curve back to E on AB. With B' as a center and radius BE strike an arc which intersects the parallel through C in E' and extend $B'E'$ to D' on the parallel through A. Now with B' as center and the radius $B'D'$ strike an arc which intersects $B'C$ in D. Then $B'D$ is the desired corrected temperature rise.

To show that this construction is valid consider the similar right triangles $FD'B'$ and $CE'B'$ in Fig. 5. Evidently $B'D'=B'E'\cdot B'F/B'C$. But $B'E'=BE=(\theta_x-\theta_0)$ and $B'F=BA=(\theta_{max}-\theta_0)$ by construction. From Eq. (42) it follows that $B'C=(\theta_{inf}-\theta_0)$. Therefore $B'D'=(\theta_x-\theta_0)\times(\theta_{max}-\theta_0)/(\theta_{inf}-\theta_0)$; and since $B'D=B'D'$ it follows that $B'D=(\theta_x^*-\theta_0^*)$ from Eq. (47).

We record here our pleasure in receiving several suggestions from members of both the chemistry and physics staffs. We are particularly grateful to Professor O. M. Arnold for his many excellent criticisms. In conclusion, we express here our appreciation for the kind assistance of Miss Harriet Peck, librarian at Rensselaer Polytechnic Institute.

[33] See ASTM Standards D271–33, p. 423 (1936).

Kinetics of Degradation and Size Distribution of Long Chain Polymers*

ROBERT SIMHA†
Department of Chemistry, Columbia University, New York, New York

(Received March 7, 1941)

The rate equations for the depolymerization process of linear chain molecules are formulated and a general solution is given. Three special cases are considered: (a) Equal disintegration probability for all linkages independently of their position in the chain. The results confirm and extend the ones previously found from statistical considerations. (b) Preferred breaking at the ends. (c) Equal disintegration probability for all chains independently of their size. Equations for the decrease of number and weight average molecular weight with time are given in each case.

IN several publications the problem of the size distribution of degradation products, generated by the splitting of long chains, has been considered.[1] By a treatment, based on statistical considerations, an expression for the distribution of molecular weights was found as function of the molecular weight of the monomer, of the initial chain length or distribution of chain lengths and of the average degree of depolymerization, i.e., the average number of bonds cut per original molecule divided by the total number of bonds in the original molecule.

As pointed out at previous occasions, a complete treatment of the kinetics of degradation has two aspects: first a determination of the distribution of all possible chain lengths in the reaction batch at different stages of the reaction; second, an investigation of the change of the properties of the system, like average molecular weight, with time. A complete solution contains implicitly the answer to both questions.

Both problems have been considered by Kuhn and Klages, respectively, for some special cases. However the results are valid only for random splitting of infinitely long chains. Montroll and Simha have given a complete statistical solution for random breaking.[2] This assumption seems reasonable in many cases for large degradation products. Nevertheless it appears desirable to develop a general theory of the breaking of long chains comprising both aspects of the problem. In case of non-random splitting a statistical calculation is difficult. It becomes more promising, and of course necessary for a determination of the time dependence, to start from the rate equations for each degrading species. The solution of these equations gives the decrease of the amount of each component present and therefore also the decrease of average molecular weight with time.

MATHEMATICAL DEVELOPMENT

Let us assume that at the time $t=0$ there are $N_j(0)$ molecules consisting of j monomers and having $j-1$ links, where $n \geqq j \geqq 2$. These molecules are exposed to a depolymerization process, e.g., hydrolysis or thermal decomposition, characterized by a set of rate constants $k_i^{(j)}$, where the lower index gives the ordinal number of the respective bond and the upper one the degree of polymerization of the corresponding molecule. Then we have the following set of differential equations:

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

† Lalor Foundation Fellow 1940–41.

[1] W. Kuhn, Ber. **63**, 1503 (1930); Zeits. f. physik. Chemie **A159**, 368 (1932); F. Klages, *ibid.* **A159**, 357 (1932); W. H. Durfee and Z. I. Kertesz, J. Am. Chem. Soc. **62**, 1196 (1940); H. Mark and R. Simha, Trans. Faraday Soc. **36**, 611 (1940); E. W. Montroll and R. Simha, J. Chem. Phys. **8**, 721 (1940). Also I. Sakurada and S. Okamura, Zeits. f. physik. Chemie **A187**, 289 (1940).

[2] The same formalism has been applied to a treatment of the removal of substituents from mixed vinyl polymers; compare F. T. Wall, J. Am. Chem. Soc. **62**, 803 (1940); R. Simha, *ibid.* **63**, 1479 (1941).

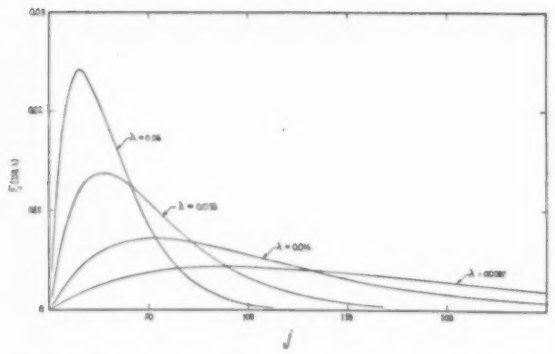


FIG. 1. Fractions of monomers involved in j-mers for various values of the reduced time, $\lambda = kt$, and for initial degree of polymerization, $n = 350$, according to Eqs. (2a′, 1) and (3′).

$$
\left.
\begin{aligned}
&\frac{dN_n}{dt} = -\sum_{i=1}^{n-1} k_i^{(n)} \cdot N_n, \\
&\frac{dN_{n-1}}{dt} = (k_1^{(n)} + k_{n-1}^{(n)}) N_n - \sum_{i=1}^{n-2} k_i^{(n-1)} \cdot N_{n-1}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\frac{dN_j}{dt} = (k_{n-j}^{(n)} + k_j^{(n)}) N_n + \\
&\qquad + (k_{n-j-1}^{(n-1)} + k_j^{(n-1)}) N_{n-1} + \cdots \\
&\qquad + \cdots (k_1^{(j+1)} + k_j^{(j+1)}) N_{j+1} \\
&\qquad - \sum_{i=1}^{j-1} k_i^{(j)} \cdot N_j, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\frac{dN_2}{dt} = (k_{n-2}^{(n)} + k_2^{(n)}) N_n + (k_2^{(n-1)} + k_{n-3}^{(n-1)}) N_{n-1} \\
&\qquad + \cdots (k_1^{(3)} + k_2^{(3)}) N_3 - k_1^{(2)} N_2.
\end{aligned}
\right\} (1)
$$

These equations express the fact that a particular chain can be generated by disintegration of any of the longer ones and is in turn annihilated by decomposition. They are somewhat similar to those describing the process of radioactive disintegration treated by Bateman, Ruark, Devol and others.[3] Recently Delbrück[4] has considered fluctuations in autocatalytic reactions and has obtained similar equations. System (1) however is more complicated, due to the fact that a particular species can be created by all the preceding ones. If both ends of a chain molecule are equal, then because of reasons of symmetry $k_i^{(j)} = k_{j-i}^{(j)}$. Having solved (1), we find the number of monomers from the condition that the sum of free and bound monomers is constant throughout the reaction process. Finally we calculate the number and weight average molecular weight defined by Kraemer and Lansing[5] from the equations:

$$\frac{M_n}{M_0} = \frac{\sum_j jN_j}{\sum_j N_j}; \quad \frac{M_w}{M_0} = \frac{\sum_j j^2 N_j}{\sum_j jN_j},$$

where M_0 denotes the molecular weight of the repeating unit.

We write the solution of (1) in the form;

$$N_j = \sum_{l=1}^{n} h_l a_{jl} \exp(\lambda_l t).$$

Solving the secular determinant for the λ_l we find

$$\lambda_1 = 0, \quad \lambda_l = -\sum_{i=1}^{l-1} k_i^{(l)} \text{ for } l = 2, 3 \cdots n.$$

The h_l are constants determined by the initial conditions. The a_{jl} are solutions of the following system of linear equations:

$$
\begin{aligned}
&-\lambda_l a_{1l} + 2k_1^{(2)} a_{2l} + \cdots 2k_1^{(j)} a_{jl} + \cdots 2k_1^{(n)} a_{nl} = 0, \\
&(-k_1^{(2)} - \lambda_l) a_{2l} + \cdots 2k_2^{(j)} a_{jl} + \cdots 2k_2^{(n)} a_{nl} = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\qquad\qquad (-\sum_{i=1}^{n-1} k_i^{(n)} - \lambda_l) a_{nl} = 0.
\end{aligned}
$$

If the above symmetry condition is not fulfilled we must write $k_j^{(n)} + k_{n-j}^{(n)}$ instead of $2k_j^{(n)}$. We see from these equations that $a_{jl} = 0$ for $j > l$. The remaining a_{jl} obey the simpler relationships:

$$
\left.
\begin{aligned}
&-\lambda_1 a_{1l} + 2k_1^{(2)} a_{2l} + \cdots 2k_1^{(l)} a_{ll} = 0, \\
&(-k_1^{(2)} - \lambda_l) a_{2l} + \cdots 2k_2^{(l)} a_{ll} = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&\qquad (-\sum_{i=1}^{l-2} k_i^{(l-1)} - \lambda_l) a_{l-1l} + 2k_{l-1}^{(l)} a_{ll} = 0.
\end{aligned}
\right\} (2)
$$

In this way we find by successive solution the coefficients a_{jl} for $j < l$ in terms of the rate constants and of the quantities a_{ll}. The final

[3] H. Bateman, Proc. Camb. Phil. Soc. **15**, 423 (1910). A. E. Ruark and L. Devol, Phys. Rev. **49**, 355 (1936).

[4] M. Delbrück, J. Chem. Phys. **8**, 120 (1940). The theory of polymerization reactions leads to similar equations if the process can be described as an addition of monomers; compare H. Dostal and H. Mark, Zeits. f. physik. Chemie **B29**, 299 (1935).

[5] E. O. Kraemer and W. D. Lansing, J. Phys. Chem. **39**, 153 (1935).

result for the $N_j(t)$ appears in the form:

$$\left.\begin{aligned} N_j(t) &= h_j a_{jj} \exp\left[-\sum_{i=1}^{j-1} k_i^{(j)} t\right] + \\ &\quad + h_{j+1} a_{jj+1} \exp\left[-\sum_{i=1}^{j} k_i^{(j+1)} t\right] + \\ &\quad + \cdots h_n a_{jn} \exp\left[-\sum_{i=1}^{n-1} k_i^{(n)} t\right] \\ N_1(t) &= \sum_{i=1}^{n} i N_i(0) - \sum_{i=2}^{n} i N_i(t) \end{aligned}\right\} \quad (2a)$$

if we denote by $N_i(0)$ the number of i-mers initially present. Equations (2) and (2a) represent the most general solution of the rate equations (1), except for the case when the solutions of the secular determinant are degenerate. We shall encounter such a degeneracy in the following discussion of special cases.

Discussion

1. Let us assume that all the linkages are broken at random. In this case $k_i^{(j)} = k$, independent of i and j and the rate equations reduce to:

$$\frac{dN_j}{dt} = 2k \sum_{i=j+1}^{n} N_i - k(j-1)N_j. \qquad (1, 1)$$

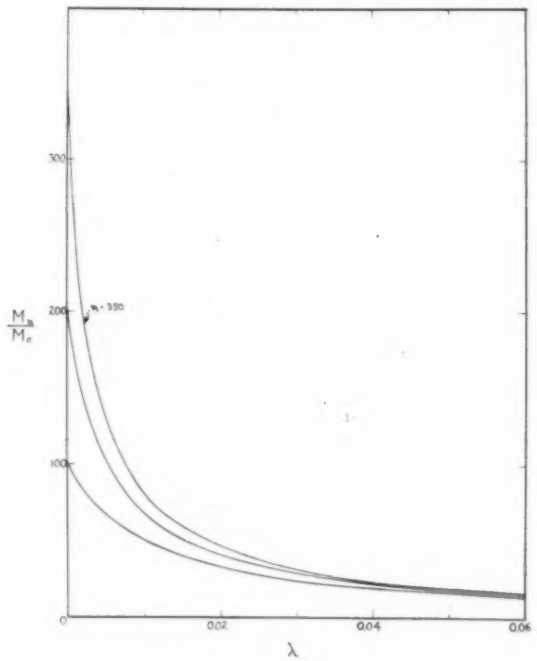


Fig. 2. Variation of number average molecular weight with reduced time, λ, for various initial chain lengths, according to Eq. (4a).

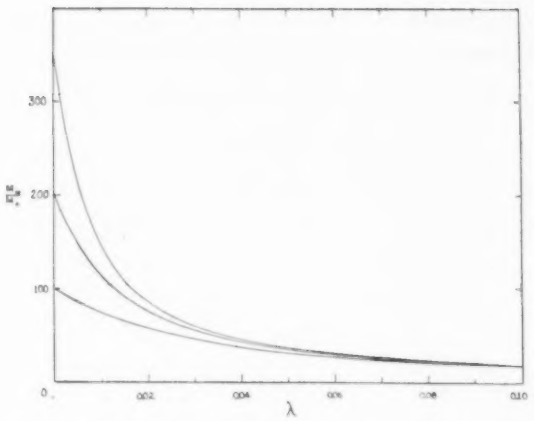


Fig. 3. Variation of weight average molecular weight with reduced time, λ, for various initial chain lengths, according to Eq. (4b).

From (2) and (2a) we find the solution

$$N_j(t) = e^{-k(j-1)t}\left[h_j a_{jj} - 2h_{j+1}a_{j+1j+1}e^{-kt} + h_{j+2}a_{j+2j+2}e^{-2kt}\right] \quad (2a, 1)$$

for $n \geqq j \geqq 2$. If the material was at $t=0$ ideally homogenous, then $N_j(0)=0$ for $j \neq n$. For the integration constants we find

$$h_j a_{jj} = (n-j+1)N_n(0).$$

Introducing

$$\alpha = 1 - e^{-kt}, \qquad (3')$$

we can write the solution as:

$$\left.\begin{aligned} N_j &= N_n(0)(1-\alpha)^{j-1}\alpha[2+(n-j-1)\alpha]; \\ &\qquad 1 \leqq j \leqq n-1 \\ N_n &= N_n(0)(1-\alpha)^{n-1}. \end{aligned}\right\} \quad (2a', 1)$$

This result is identical with that previously found from statistical considerations by Montroll and the author. In their work α denoted the average degree of depolymerization as defined above. Here α is expressed as function of the time elapsed since the start of the reaction. This time dependence has been previously derived by Kuhn[1] from the condition that the rate of decrease of the number of linkages is proportional to the number of bonds present. That is, in our notation:

$$-\frac{d}{dt}\sum_{i=1}^{n}(i-1)N_i = k\sum_{i=1}^{n}(i-1)N_i.$$

This equation follows immediately from our rate

equations (1a). On the other hand, the degree of depolymerization is defined as the ratio:

$$\alpha=\frac{\sum_{i=1}^{n}(i-1)N_i(0)-\sum_{i=1}^{n}(i-1)N_i(t)}{\sum_{i=1}^{n}(i-1)N_i(0)}$$

$$=\frac{\sum_{i=1}^{n}N_i(t)-\sum_{i=1}^{n}N_i(0)}{\sum_{i=1}^{n}(i-1)N_i(0)}=\frac{\sum_{i=1}^{n}N_i(t)-N_n(0)}{(n-1)N_n(0)}, \quad (3)$$

where the second identity holds for an initially homogenous substance. Inserting from (2a, 1) and considering the boundary conditions we find after a simple calculation the above result for α. This follows also from the rate equation for the breaking of linkages. We have in this way rederived the statistical results, and furthermore found the time dependence of the inhomogeneity of the degrading system.[6] Figures 1, 2, and 3, give the weight fraction distribution, the number average molecular weight M_n and the weight average molecular weight M_w as function of the reduced time $\lambda=kt$. The problem of the size distribution produced by the uniform splitting of long chains appears in this way completely solved.

The experiments of Wolfrom and others[7] on the hydrolysis of potato starch and of methylated cellulose and those of Flory[8] on the degradation of polyesters by alcohols indicate such a mechanism to be present. However no fractionation of the reaction products and subsequent analysis of the distribution curve was undertaken.

Inserting the expression (3′) into the equation previously found by Montroll and the author[1] for the number average molecular weight M_n, namely,

$$M_n=\frac{n}{1+\alpha(n-1)}, \quad (4a)$$

we find the equation previously derived by Wolfrom, Sowden and Lassettre[7]

$$\left(\frac{M_n}{M_0}-1\right)\Big/\frac{M_n}{M_0}=\frac{n-1}{n}e^{-kt}.$$

The decrease of the weight average molecular weight with time is found from (3′) and the following equation derived by Montroll and Simha:[1]

$$\frac{M_w}{M_0}=\frac{n\alpha^2+2(1-\alpha)[(1-\alpha)^n+n\alpha-1]}{n\alpha^2}. \quad (4b)$$

2. The other extreme case is that of preferred breaking at the ends, i.e., the probability for breaking end bonds is so much larger than the corresponding rates for the interior of the chains that one has during the whole cracking process only very large components and monomeric material in the mixture, without any appreciable amount of intermediate sizes. There are indications according to experiments of Mark and Raff[9] that the thermal depolymerization of polystyrene in liquid phase at about 350° proceeds in this way. As a first approximation we assume that $k_1^{(j)}=k_{j-1}^{(j)}=k^{(j)}$, while all the other rate constants are zero. The differential equations then reduce to

$$\frac{dN_j}{dt}=2k^{(j+1)}N_{j+1}-2k^{(j)}N_j; \quad n\geqq j\geqq 2. \quad (1, 2)$$

They are of the general type treated by Bateman[3] and have the solution

$$N_j=\sum_{i=j}^{n}c_i\exp\left[-2k^{(i)}t\right]$$

with

$$c_i=N_n(0)\times\frac{k^{(n)}k^{(n-1)}\cdots k^{(j+1)}}{(k^{(n)}-k^{(i)})(k^{(n-1)}-k^{(i)})\cdots(k^{(j)}-k^{(i)})}. \quad (2a, 2)$$

This result holds also for $j=2$ if we denote the rate constant for the dimer by $2k^{(2)}$. A separate treatment of the case of equal disintegration probabilities, $k^{(i)}=k^{(j)}=k^{(2)}=k$ gives the solution:

$$\left.\begin{aligned}
&N_j=N_n(0)e^{-z}\cdot\frac{z^{n-j}}{(n-j)!}; \quad n\geqq j>2\\
&N_2=N_n(0)e^{-z}\cdot\frac{z^{n-2}}{(n-2)!}+N_n(0)R\\
&N_1=nN_n(0)-N_n(0)e^{-z}\left\{n\cdot\frac{z^{n-2}}{(n-2)!}+\right.\\
&\left.+(n-z)\left(e^z-\sum_{r=n-2}^{\infty}\frac{z^r}{r!}\right)\right\}+2N_n(0)R
\end{aligned}\right\} \quad (2a, 2')$$

[6] The statistical result involves the trivial assumption that the N_j are very large numbers. This condition is here introduced implicitly by describing the rate process by means of differential equations rather than difference equations.

[7] M. L. Wolfrom, E. N. Lassettre, D. R. Myers and J. C. Sowden, J. Am. Chem. Soc. **61**, 1072, 2172 (1939).

[8] P. J. Flory, J. Am. Chem. Soc. **62**, 2255 (1940).

[9] H. Mark and R. Raff, private communication.

with $z=2kt$ and

$$R=e^{-z}\cdot 2^{n-2}\sum_{i=n-1}^{\infty}\left(\frac{z}{2}\right)^{i}\frac{1}{i!}.$$

Equations (2a, 2) and (2a, 2′) describe the course of the degradation process for an initially homogenous material of degree of polymerization n. For large values of n and for the first stages of the reaction (small values of z) the number of monomers produced, increases linearly with time:

$$N_1\approx N_n(0)z.$$

For the number and weight average molecular weight we find, respectively:

$$\frac{M_n}{M_0}=\frac{1}{1-\frac{e^{-z}}{n}\left\{(n-1)\frac{z^{n-2}}{(n-2)!}+(n-z-1)\left(e^{z}-\sum_{r=n-2}^{\infty}\frac{z^r}{r!}\right)\right\}-\frac{R}{n}}, \tag{4a′}$$

$$\frac{M_w}{M_0}=1+e^{-z}\left\{(n-1)\frac{z^{n-2}}{(n-2)!}+\left(n-2z-1+\frac{2z}{n}\right)\frac{z^{n-3}}{(n-3)!}\right.$$
$$\left.+\left(1-\frac{z}{n}\right)\left(n-z-1+\frac{z}{n-z}\right)\left(e^{z}-\sum_{n-3}^{\infty}\frac{z^r}{r!}\right)\right\}+\frac{2R}{n}. \tag{4b′}$$

Figure 4 shows a graph of these functions for $n=300$. As to be expected, the weight average decreases very slowly because the monomers produced contribute very little to M_w, while they influence relatively strongly the decrease of M_n.

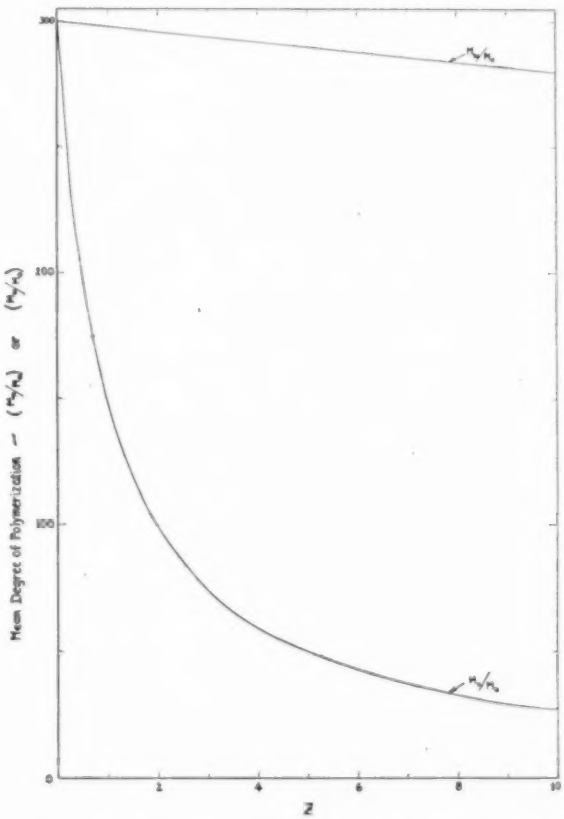


FIG. 4. Variation of number and weight average molecular weight with reduced time, $z=2kt$, for $n=300$, according to Eqs. (4a′) and (4b′).

3. In our first case the disintegration probability for a chain was proportional to its size or more correctly to the number of linkages by which it is formed. Let us now consider a case where the chance of decomposing is roughly equal for all molecules independently of their size. The probability for a particular bond in a particular chain being cut is then inversely proportional to the number of linkages in that chain. To the author's knowledge no experimental data are available as yet, but such an assumption might give an approximate description of the catalyzed cracking of long chain molecules like polystyrene.[10] It seems possible that surface catalysis affects each molecule approximately in the same way, at least within certain size limits, and thus justifies the above assumption. Our rate equations then read as follows:

$$\frac{dN_j}{dt}=2k\sum_{i=j}^{n-1}\frac{N_{i+1}}{i}-kN_j;\quad n\geqq j\geqq 2. \tag{1, 3}$$

The first step, the formation of a molecule j by disintegration of an i-mer has a rate constant $2k/i-1$ because there are two possibilities for this step. k is the rate constant for the decomposition of chains of any length. The previous solution (2) and (2a) does not apply here because of

[10] I am indebted to Professor H. Mark with whom I had the opportunity to discuss this point.

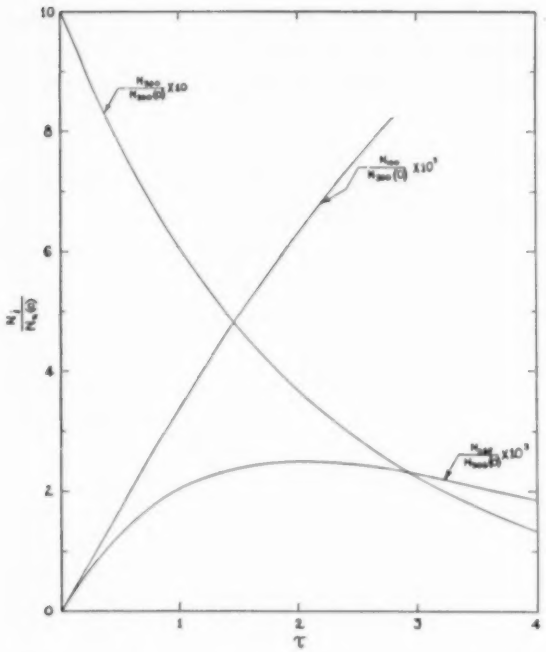


FIG. 5. Variation of fractional number of chains, $N_j/N_n(0)$, with reduced time, $\tau=2kt$, according to Eq. (2a, 3).

the degeneracy of the determinant. Subtracting the respective equations for N_j and N_{j+1} from each other we get

$$\frac{dN_j}{dt}+kN_j=\frac{dN_{j+1}}{dt}\quad+kN_{j+1}\left(1+\frac{2}{j}\right)=F_{j+1}.$$

Therefrom

$$N_j=e^{-kt}\left[C_j+\int F_{j+1}e^{kt}dt\right].$$

C_j is an integration constant. If again for $t=0$, $N_n=N_n(0)$ and all the other N_j are zero, C_j vanishes. By successive integration we find with $\tau=2kt$

$$F_{j+1}=\frac{N_n(0)}{n-1}e^{-\tau/2}\cdot 2k\left\{1+\tau S_1^{(j+1)}+\frac{\tau^2}{2!}S_2^{(j+1)}+\cdots+\frac{\tau^{n-j-1}}{(n-j-1)!}S_{n-j-1}^{(j+1)}\right\},$$

$$N_j=\frac{N_n(0)}{n-1}e^{-\tau/2}\left\{\tau+\frac{\tau^2}{2!}S_1^{(j+1)}+\cdots\frac{\tau^{n-j}}{(n-j)!}S_{n-j-1}^{(j+1)}\right\}\qquad(2a,\ 3)$$

$$N_n=N_n(0)e^{-\tau/2}.$$

The coefficients $S_k^{(j+1)}$ are defined as follows:

$$S_0^{(j+1)}=1,$$

$$S_k^{(j+1)}=\sum_{\nu_1=j+k-1}^{n-2}\ \sum_{\nu_2=j+k-2}^{\nu_1-1}\cdots\sum_{\nu_k=j}^{\nu_{k-1}-1}\frac{1}{\nu_1}\frac{1}{\nu_2}\cdots\frac{1}{\nu_k}.\qquad(5)$$

From the relationship

$$F_{j+1}-F_j=-\frac{2k}{j-1}\cdot N_j,$$

we obtain a recurrence formula for the $S_k^{(j+1)}$

$$S_{k+1}^{(j+1)}+\frac{1}{j-1}S_k^{(j+1)}=S_{k+1}^{(j)};\quad k\leqq n-j-2$$

$$\frac{1}{j-1}S_{n-j-1}^{(j+1)}=S_{n-j}^{(j)}.$$

The values for the upper and lower indices are limited by the condition $k+j+1\leqq n$. It is possible to express these coefficients in terms of the ψ function. For integer values of the argument it is defined by

$$\psi(n)=-C+1+\frac{1}{2}+\cdots\frac{1}{n};$$

$$\psi(n+1)=\psi(n)+\frac{1}{n+1}.$$

C is the Euler-Mascheroni constant and has the value 0.57722. For sufficiently large values of the argument we have the expansion:[11]

$$\psi(n)\approx\ln n+\frac{1}{2n}-\frac{1}{12}\frac{1}{n^2}+\frac{1}{120}\frac{1}{n^4}-\frac{1}{252}\frac{1}{n^6}+0\left(\frac{1}{n^8}\right).$$

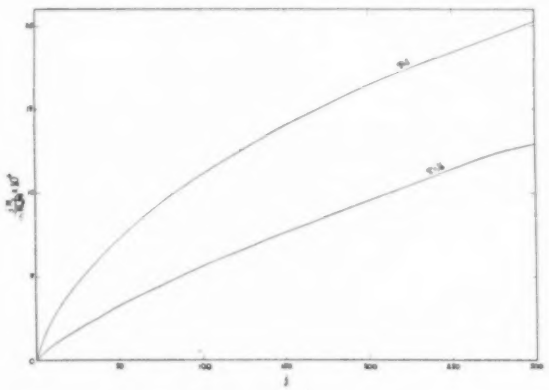


FIG. 6. Fraction of monomers involved in j-mers for various values of the reduced time, τ, and for $n=300$, according to Eq. (2a, 3).

[11] E. Madelung, *Die Mathematishen Hilfsmittel des Physikers* (J. Springer, Berlin, 1936).

We shall make use of this series in the numerical calculation of the average molecular weight.

Now obviously

$$S_1^{(j+1)}=\psi(n-2)-\psi(j-1).$$

From the equations derived in the Appendix we conclude

$$S_2^{(j+1)}=\frac{1}{2}\Big\{\psi(n-2)\psi(n-3)+\frac{\psi(n-2)}{n-2}$$
$$-2\psi(n-2)\psi(j-1)+\psi^2(j)-\frac{2\psi(j)}{j}$$
$$+\sum_{k=1}^{j-1}\frac{1}{(k+1)^2}-\sum_{k=1}^{n-3}\frac{1}{(k+1)^2}\Big\}. \quad (5a)$$

The use of the infinite series

$$\sum_{k=1}^{\infty}\frac{1}{(k+1)^2}=\frac{\pi^2}{6}-1$$

facilitates the evaluation of the two sums in (5a) or at least of the second one. In the range of n values in which we are interested here, the value of the second sum is practically equal to that of the series. For intermediate values of j correction terms, starting from the series, can be calculated for any desired degree of accuracy. It is possible to compute higher terms in the expression (2a, 3) in a way analogous to that used in the Appendix for $S_2^{(j+1)}$. The computations however soon become very cumbersome. Therefore, at present, a more accurate evaluation of the higher terms has been omitted. They decrease quite rapidly with their ordinal number j, as can also be seen from the following numerical example. Therefore even for values of τ of the order of magnitude of one the accuracy obtainable in this way for N_j is sufficient for intermediate sizes, e.g., $j=100$, if $n=300$. A rough estimation of the coefficients can be made by using the first term of the expansion for $\psi(n)$. If

$$\frac{k+1}{(j-1)[\ln(n-k-2)-\ln(j-1)]}\ll 1$$

and $\ln(n-k-2)\approx\ln(n-k-1)$, then

$$S_k^{(j+1)}\approx\frac{1}{k!}[\ln(n-k-1)-\ln(j-1)]^k.$$

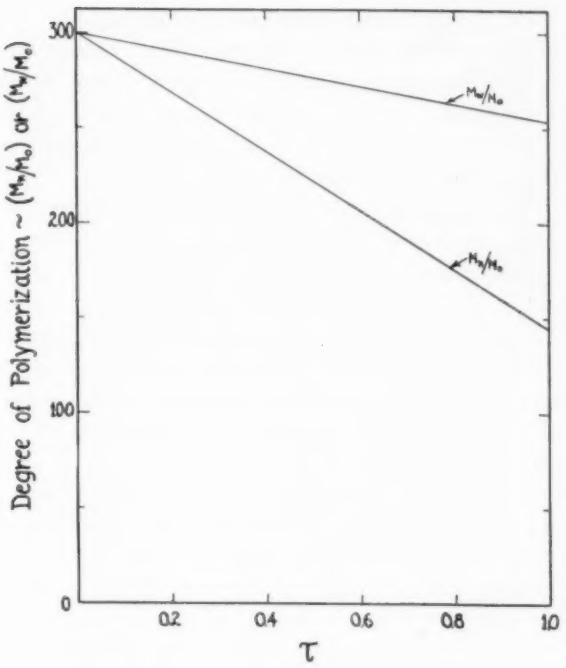


FIG. 7. Variation of number and weight average molecular weight with reduced time, τ, for $n=300$, according to equations derived in the Appendix.

This equation can be used when k is small; j must be large enough to allow the substitution of the logarithm for the function. For $j=11$ the corresponding error is about $2\frac{1}{2}$ percent. $j-1$, however, must be small compared with $n-k-2$.

For the computation of N_1, M_w and M_n we need the sums

$$\sum_{j=2}^{n}N_j,\quad\sum_{j=2}^{n}jN_j,\quad\sum_{j=2}^{n}j^2N_j.$$

We get

$$\sum_{j=2}^{n}N_j=N_n(0)e^{-\tau/2}+\frac{N_n(0)}{n-1}e^{-\tau/2}$$
$$\times\Big\{\tau(n-2)+\frac{\tau^2}{2!}\sum_{j=2}^{n-2}S_1^{(j+1)}+\cdots+\cdots\frac{\tau^{k+1}}{(k+1)!}$$
$$\times\sum_{j=2}^{n-k-1}S_k^{(j+1)}+\cdots\frac{\tau^{n-2}}{(n-2)!}S_{n-3}^{(3)}\Big\}$$

and similar results for the other sums. Figure 5 gives the decrease of the number of various species with reduced time. Figure 6 shows the

total weight distribution during the first stages of the reaction. The final graph, Fig. 7, reproduces the decrease of the mean degree of polymerization with time. As is to be expected, the curves are much steeper for random splitting of bonds than for random splitting of chains.

The above equations have all been formulated explicitly for an initially homogenous material. If we have a whole distribution then the final result for $N_j(t)$ is a superposition of partial solutions of the type written down, each containing a factor $N_k(0)$ with $n \geqq k \geqq j$. If we represent the initial conditions by a smooth distribution curve we can write an integral between the limits n and j instead.

Theoretical attempts to obtain the disintegration constant of a chain molecule as function of the position of the bond in the chain, its strength and the size of the chain hitherto have not been very successful. Pelzer[12] has examined the way in which an isolated chain molecule can distribute excess energy among its linkages. If all the masses and bonds are identical, equal probabilities for end and middle bonds and no dependence upon the length of the chain is found. According to Schmidt,[13] a different mechanism is operative in the cracking, e.g., of hydrocarbons. The course of the reaction in the hydrolysis of cellulose can be best described, according to Freudenberg and Blomquist,[14] by the assumption that end bonds react twice as fast as those in the interior of the molecule.

It is very desirable to have more experimental data on depolymerization reactions. The change of the mean molecular weight with time as well as the whole distribution curve at different stages of the reaction should be determined. From the point of view of the theory it appears most suitable to study thermal cracking processes, because in this way the conditions imposed by the theory could be best verified. Such investigations might also represent a first step to a study of equilibria between polymerization and depolymerization processes.[15]

The author wishes to express his thanks to the Lalor Foundation whose grant of a Fellowship made this work possible.

Appendix

(a) Calculation of $S_2^{(j+1)}$

$$S_2^{(j+1)} = \sum_{\nu_1=j+1}^{n-2} \frac{\psi(\nu_1-1)}{\nu_1} - \psi(j-1)[\psi(n-2)-\psi(j)].$$

It can be shown that

$$\sum_{k=1}^{n} \frac{\psi(k)}{k+1} = \psi(n+1)\psi(n) + C\psi(1) - \sum_{k=1}^{n} \frac{\psi(k)}{k}.$$

Proof:

$$\begin{aligned}\sum_{k=1}^{n} \frac{\psi(k)}{k+1} &= \frac{1}{2}(-C+1)+\frac{1}{3}\left(-C+1+\frac{1}{2}+\frac{1}{3}\right)+\cdots\frac{1}{n+1}\left(-C+1+\frac{1}{2}+\cdots\frac{1}{n}\right)\\ &= -C[\psi(n+1)-\psi(1)]+[\psi(n+1)-\psi(1)]+\cdots\frac{1}{n}[\psi(n+1)-\psi(n)]\\ &= \psi(n+1)\psi(n)+C\psi(1)-\sum_{k=1}^{n}\frac{\psi(k)}{k}.\end{aligned}$$

On the other hand, using the recurrence formula for $\psi(n)$,

$$\sum_{k=1}^{n} \frac{\psi(k)}{k} = \sum_{k=1}^{n} \frac{\psi(k+1)}{k+1} - \frac{\psi(n+1)}{n+1} + \psi(1) = \sum_{k=1}^{n} \frac{\psi(k)}{k+1} + \sum_{k=1}^{n} \frac{1}{(k+1)^2} - \frac{\psi(n+1)}{n+1} + \psi(1).$$

[12] H. Pelzer, Zeits. f. Elektrochemie **39**, 608 (1933).
[13] O. Schmidt, Zeits. f. physik. Chemie **A159**, 337 (1932).
[14] K. Freudenberg, and G. Blomquist, Ber. **68**, 2070 (1935). See also Alf af Ekenstam. *ibid.* **69**, 549, 553 (1936).
[15] The resulting size distribution has been derived in terms of an equilibrium constant under simplifying assumptions by G. V. Schulz, Zeits. f. physik. Chemie **B43**, 25 (1939).

Making use of these expressions and taking the proper limits we find Eq. (5a) given in the paper.

(b) Calculation of $\sum_{j=2}^{n} j^i N_j$, $i=0, 1, 2$.

$$\sum_{j=2}^{n-2} S_1^{(j+1)} = (n-3)\psi(n-2) - \sum_{j=2}^{n-2} \psi(j-1) = (n-3)\psi(n-2) - \sum_{j=1}^{n-3} \psi(j),$$

$$\sum_{j=1}^{n-3} \psi(j) = -C(n-3) + (n-3) + \frac{1}{2}(n-4) + \cdots \frac{1}{n-3} = 3C + n\psi(n-3) - \sum_{k=1}^{n-3} \frac{k+2}{k} = C + (n-2)\psi(n-3) - (n-3);$$

$$\sum_{j=2}^{n-3} S_1^{(j+1)} = (n-2) - \psi(n-2) - C.$$

The calculation of the sum

$$\sum_{j=2}^{n-3} S_2^{(j+1)}$$

requires an evaluation of the following partial sums:

$$\sum_{j=2}^{n-3} \frac{\psi(j)}{j} = \frac{1}{2}\left[(C-1)\psi(1) + \psi(n-2)\psi(n-3) - \frac{\psi(n-2)}{n-2} + \sum_{k=1}^{n-3} \frac{1}{(k+1)^2}\right]$$

as can be seen from the results derived under (a), and

$$\sum_{j=2}^{n-3} \sum_{k=1}^{j-1} \frac{1}{(k+1)^2} = (n-2) \sum_{k=1}^{n-4} \frac{1}{(k+1)^2} - \psi(n-3) + \psi(1)$$

obtained by direct computation. The final expression is:

$$\sum_{j=2}^{n-3} S_2^{(j+1)} = \frac{1}{2}\left\{\sum_{j=2}^{n-3} \psi^2(j) + \frac{\psi(n-2)}{n-2}(n-3) - 2(n-3)\psi(n-2)\psi(n-4) + (n-5)\psi(n-2)\psi(n-3)\right.$$

$$\left. + 2(n-4)\psi(n-2) - 2C\psi(n-2) + (2-C)\psi(1) - \psi(n-3) - \frac{(n-3)}{(n-2)^2} + \sum_{k=1}^{n-4} \frac{1}{(k+1)^2}\right\}.$$

In the numerical examples the sum

$$\sum_{j=2}^{n-3} \psi^2(j)$$

has been evaluated with the aid of Euler's summation formula,[11] valid if the function and its derivatives appearing, are continuous:

$$\sum_{k=0}^{m} f(k) = \int_0^m f(t)dt + \frac{1}{2}[f(0) + f(m)] + \sum_{k=1}^{n} \frac{B_{2k}}{(2k)!}\{f^{(2k-1)}(m) - f^{(2k-1)}(0)\} + m\frac{B_{2n+2}}{(2n+2)!} f^{(2n+2)}(\vartheta m)$$

with $m, n = 1, 2\cdots$; $0 < \vartheta < 1$. The B_{2k} are Bernoulli's numbers, $B_0 = 1$, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$. In our applications it is sufficient to take two terms in the last sum of Euler's formula. In the calculation of this and the following integrals the expansion for the ψ function up to terms $1/t^4$ was used. The numerical values of $\psi(t)$ and its derivatives are given in Davies, *Tables of Functions*. We obtain

$$\sum_{j=2}^{n} jN_j \approx N_n(0)e^{-\tau/2}\left\{n + \frac{\tau}{n-1}\left[\frac{(n-1)n}{2} - 1\right] + \frac{\tau^2}{2!(n-1)} \sum_{j=2}^{n-2} jS_1^{(j+1)} + \frac{\tau^3}{3!(n-1)} \sum_{j=2}^{n-3} jS_2^{(j+1)} + O(\tau^4)\right\}$$

and

$$\sum_{j=2}^{n} j^2 N_j \approx N_n(0)e^{-\tau/2}\left\{n^2 + \frac{\tau}{n-1}\left[\frac{(n-1)n(2n-1)}{6} - 1\right] + \frac{\tau^2}{2!(n-1)} \sum_{j=2}^{n-2} j^2 S_1^{(j+1)} + \frac{\tau^3}{3!(n-1)} \sum_{j=2}^{n-3} j^2 S_2^{(j+1)} + O(\tau^4)\right\}.$$

We consider the following sums:

$$\sum_{k=1}^{n} k\psi(k) = \frac{n(n+1)}{2}\psi(n) - \frac{1}{4}n(n-1); \quad \sum_{k=1}^{n} k^2\psi(k) = \frac{n(n+1)(2n+1)}{6}\left[\psi(n) - \frac{1}{3}\right] + \frac{1}{4}n(n+1) - \frac{n}{6};$$

$$\sum_{j=2}^{n-3} j \sum_{k=1}^{j-1} \frac{1}{(k+1)^2} = \frac{(n-3)(n-2)}{2} \sum_{k=1}^{n-4} \frac{1}{(k+1)^2} + \frac{1}{2}[\psi(n-3) - \psi(1)] - \frac{(n-4)}{2};$$

$$\sum_{j=2}^{n-3} j^2 \sum_{k=1}^{j-1} \frac{1}{(k+1)^2} = \frac{(n-3)(n-2)(2n-5)}{6} \sum_{k=1}^{n-4} \frac{1}{(k+1)^2} + \frac{(n-4)}{2} - \frac{1}{3}\left[\frac{(n-3)(n-2)}{2} - 1\right] - \frac{1}{6}[\psi(n-3) - \psi(1)]$$

and the corresponding integrals

$$\int_2^{n-3} t\psi^2(t)dt \quad \text{and} \quad \int_2^{n-3} t^2\psi^2(t)dt.$$

We finally have

$$\sum_{j=2}^{n-2} jS_1^{(j+1)} = \frac{(n-3)(n+2)}{4} - \psi(n-2) + \psi(1)$$

$$\sum_{j=2}^{n-2} j^2 S_1^{(j+1)} = \frac{(n-2)(n-1)(4n+3)}{36} + \frac{(n-8)}{6} - \psi(n-2) + \psi(1).$$

For the sake of brevity, the sums over $S_2^{(j+1)}$ have not been written down explicitly, but they can easily be found with the aid of the previous expressions.

Taking in our example $n=300$ we find for instance

$$\sum_{j=2}^{n} N_j = N_n(0)e^{-\tau/2}(1+0.997\tau+0.488\tau^2+0.143\tau^3+O(\tau^4))$$

and similar equations for the other sums showing their relatively good convergence.

Consolidation Settlement of a Soil with an Impervious Top Surface

M. A. Biot and F. M. Clingan
California Institute of Technology, Pasadena, California

(Received April 8, 1941)

An investigation is made of the consolidation settlement for a soil whose top surface is impervious and loaded uniformly on an infinite strip of constant width. The settlement under the load is accompanied by considerable swelling of the unloaded area on both sides of the load.

Introduction

In a previous publication[1] the settlement of a consolidating soil under an area of constant width was obtained when the water is allowed to flow freely through the top surface both outside and under the load. We are going to investigate the same problem when the top surface is completely impervious both outside and under the load. It is clear that in this case the settlement is due to the fact that the water contained in the soil flows from under the load to unloaded regions. Since it is assumed that the water cannot escape through the top surface, this will produce a swelling of the unloaded area in the vicinity of the load. For the basic theory the reader is referred to a recent paper.[2]

I. Settlement Under a Load with Sinusoidal Distribution

In the previous paper[2] have been established the following equations for the two-dimensional problem of the consolidation of completely saturated clay with zero Poisson ratio

$$\begin{aligned} G\nabla^2 u + G\frac{\partial \epsilon}{\partial x} - \frac{\partial \sigma}{\partial x} &= 0, \\ G\nabla^2 w + G\frac{\partial \epsilon}{\partial z} - \frac{\partial \sigma}{\partial z} &= 0, \\ \nabla^2 \epsilon &= p\epsilon/c. \end{aligned} \tag{1.1}$$

In these equations u and w are the components of the displacement of the soil, $\epsilon = \partial u/\partial x + \partial w/\partial z$; G is the shear modulus for the completely consolidated clay; $c = k/a$ coefficient of consolidation; k is the coefficient of permeability; $a = 1/(2G)$ the final compressibility; σ water pressure increment in the pores; $p = \partial/\partial t$ symbolic operator. These equations must be solved with the boundary conditions (1) that all variables vanish at infinite depth $z = \infty$; (2) $\partial\sigma/\partial z = 0$ at $z = 0$; (3) $-\sigma + 2G\partial w/\partial z = -A \sin \lambda x$ at $z = 0$; (4) $\partial u/\partial z + \partial w/\partial x = 0$ at $z = 0$. The second condition expresses that no water flows from the top surface.

The last two conditions express that at the surface the normal stress is equal to the load and that the shearing stress is zero.

[1] M. A. Biot, "Consolidation settlement under a rectangular load distribution," J. App. Phys. **12**, 426 (1941).
[2] M. A. Biot, "General theory of three-dimensional consolidation," J. App. Phys. **12**, 155 (1941).

It may be verified that a solution of Eqs. (1.1) satisfying the boundary condition (1) is

$$\begin{aligned} u&=[C_1\lambda e^{-\lambda z}+C_2\lambda \exp[-(\lambda^2+p/c)^{\frac{1}{2}}z] \\ &\quad -C_3(1-\lambda z)e^{-\lambda z}]\cos\lambda x, \\ w&=[-C_1\lambda e^{-\lambda z}-C_2(\lambda^2+p/c)^{\frac{1}{2}} \\ &\quad \times\exp[-(\lambda^2+p/c)^{\frac{1}{2}}z] \\ &\quad -C_3\lambda z e^{-\lambda z}]\sin\lambda x, \\ \sigma&=2G[C_2(p/c)\exp[-(\lambda^2+p/c)^{\frac{1}{2}}z] \\ &\quad -C_3\lambda e^{-\lambda z}]\sin\lambda x, \end{aligned} \tag{1.2}$$

in which the arbitrary constants C_1, C_2 and C_3 are to be determined so that the remaining boundary conditions (2), (3) and (4) are satisfied. We find

$$\begin{aligned} (p/c)C_2(\lambda^2+p/c)^{\frac{1}{2}}-C_3\lambda^2&=0, \\ 2G\lambda^2(C_1+C_2)&=-A, \\ \lambda C_1+(\lambda^2+p/c)^{\frac{1}{2}}C_2-C_3&=0. \end{aligned} \tag{1.3}$$

We are interested in the vertical deflection w at the surface, $z=0$. This value is from (1.2) and (1.3)

$$\begin{aligned} w&=[-C_1\lambda-C_2(\lambda^2+p/c)^{\frac{1}{2}}]\sin\lambda x \\ &=-C_3\sin\lambda x. \end{aligned} \tag{1.4}$$

Solving (1.3) for C_3 we obtain

$$C_3=\frac{A}{2G\lambda}\frac{(1+p/\lambda^2c)^{\frac{1}{2}}}{(\lambda^2c/p-1)(p/\lambda^2c+1)^{\frac{1}{2}}-\lambda^2c/p}. \tag{1.5}$$

The soil deflection at the surface as a function of the time for a sudden sinusoidal load distribution is written in operational form

$$w=\frac{A\sin\lambda x}{2G\lambda}\frac{p^2/c^2\lambda^4-1-(1+p/c\lambda^2)^{\frac{1}{2}}}{p^2/c^2\lambda^4-p/c\lambda^2-1}. \tag{1.6}$$

The function represented symbolically by this equation may be calculated from the following operational expressions.

$$\frac{p^2/c^2\lambda^4}{p^2/c^2\lambda^4-p/c\lambda^2-1}1(t)=\frac{\exp[\lambda^2ct/2]}{5^{\frac{1}{2}}}\left[\sinh\left(\frac{5^{\frac{1}{2}}}{2}\lambda^2ct\right)+5^{\frac{1}{2}}\cosh\left(\frac{5^{\frac{1}{2}}}{2}\lambda^2ct\right)\right]. \tag{1.7}$$

$$\frac{1}{p^2/c^2\lambda^4-p/c\lambda^2-1}1(t)=\frac{\exp[\lambda^2ct/2]}{5^{\frac{1}{2}}}\left[-\sinh\left(\frac{5^{\frac{1}{2}}}{2}\lambda^2ct\right)+5^{\frac{1}{2}}\cosh\left(\frac{5^{\frac{1}{2}}}{2}\lambda^2ct\right)\right]-1, \tag{1.8}$$

$$\begin{aligned} \frac{(1+p/c\lambda^2)^{\frac{1}{2}}}{p^2/c^2\lambda^4-p/c\lambda^2-1}1(t)&=\frac{\exp[(1+5^{\frac{1}{2}})\lambda^2ct/2]}{5^{\frac{1}{2}}}P[(1+5^{\frac{1}{2}})\lambda(ct)^{\frac{1}{2}}]-P[\lambda(ct)^{\frac{1}{2}}] \\ &\quad -\frac{\exp[(1-5^{\frac{1}{2}})\lambda^2ct/2]}{5^{\frac{1}{2}}}P[(1-5^{\frac{1}{2}})\lambda(ct)^{\frac{1}{2}}]. \end{aligned} \tag{1.9}$$

The first two operators are tabulated[3] while the third is derived by separation into partial fractions and applying the shifting formula[4] to the well-known operator

$$p^{\frac{1}{2}}1(t)=\frac{1}{(\pi t)^{\frac{1}{2}}}. \tag{1.10}$$

With these results the deflection at the soil surface is written

$$\begin{aligned} w=\frac{A\sin\lambda x}{2G\lambda}\Big\{&1+P[\lambda(ct)^{\frac{1}{2}}] \\ &+\frac{\exp[(1+5^{\frac{1}{2}})\lambda^2ct/2]}{5^{\frac{1}{2}}}(1-P[(1+5^{\frac{1}{2}})\lambda(ct)^{\frac{1}{2}}]) \\ &-\frac{\exp[(1-5^{\frac{1}{2}})\lambda^2ct/2]}{5^{\frac{1}{2}}}(1-P[(1-5^{\frac{1}{2}})\lambda(ct)^{\frac{1}{2}}])\Big\}. \end{aligned} \tag{1.11}$$

In these formulas we have

$$P(x)=\frac{2}{\pi^{\frac{1}{2}}}\int_0^x\exp(-\xi^2)d\xi$$

[3] E. Berg, *Heaviside's Operational Calculus* (McGraw-Hill Book Company, Inc., New York, 1929), p. 165.

[4] V. Bush, *Operational Circuit Analysis* (John Wiley and Sons, New York, 1929), pp. 130, 191.

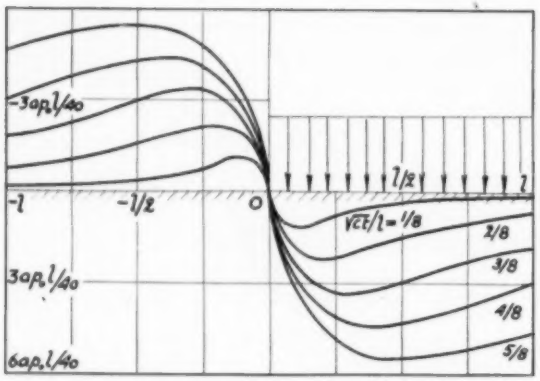


FIG. 1. Settlement of soil surface at various instants of time for a uniform load p_0 extending from $x=0$ to $x=\infty$.

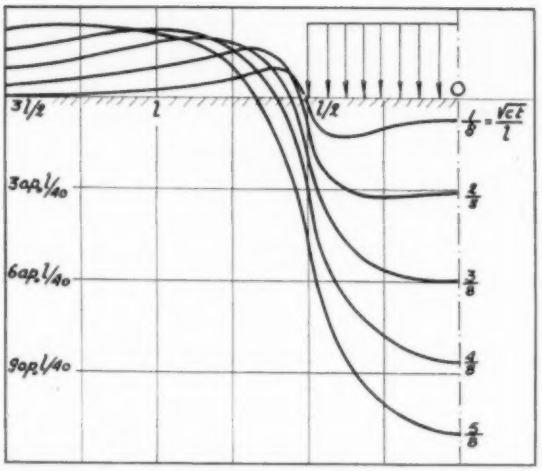


FIG. 2. Settlement of soil surface at various instants of time for a uniform load p_0 extending from $x=-l/2$ to $x=+l/2$.

a tabulated function called the probability integral.

The settlement is composed of two parts: a purely elastic deflection which occurs at the instant of application of the load and a settlement due to consolidation which occurs gradually thereafter. Since we are interested only in the latter we proceed as follows:

For $t=0$ the deflection is purely elastic, its value is

$$w_i = A \sin \lambda x / 2G\lambda. \tag{1.12}$$

The settlement due to consolidation is

$$w_s = w - w_i. \tag{1.13}$$

For convenience we will write Eq. (1.13) in the form

$$w_s = (aA/\lambda) F[\lambda(ct)^{\frac{1}{2}}] \sin \lambda x. \tag{1.14}$$

II. Settlement Under a Uniform Load with Discontinuity

For practical purposes we are interested in the differential settlement occurring when the load distribution has a discontinuity. Such a load distribution will be represented by the following function of x

$$\varphi(x) = \frac{p_0}{\pi} \int_0^\infty \frac{\sin \lambda x}{\lambda} d\lambda \tag{2.1}$$

in which case $\varphi(x) = -p_0/2$ for $x<0$, $p_0/2$ for $x>0$ with a discontinuity at $x=0$.

The settlement due to this load is obtained from (1.14) by the principle of superposition;

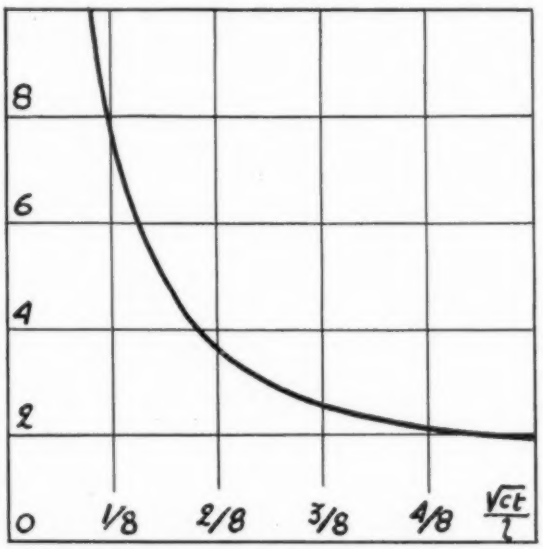


FIG. 3. Rates of the settlement at the center of the loaded area in Fig. 2 when the top surface is pervious to the settlement, when the top surface is impervious.

we find

$$w_s = \frac{ap_0}{\pi} \int_0^\infty \frac{F[\lambda(ct)^{\frac{1}{2}}]}{\lambda^2} \sin \lambda x d\lambda, \tag{2.2}$$

which can be written in the form

$$\frac{w_s}{2ap_0} \left(\frac{\pi}{ct}\right)^{\frac{1}{2}} = \frac{1}{2\pi^{\frac{1}{2}}} \int_0^\infty \frac{F(\gamma)}{\gamma^2} \sin(\gamma\xi) d\gamma \tag{2.3}$$

with $\xi = x/(ct)^{\frac{1}{2}}$.

This integral can be evaluated numerically by the use of an approximate analytical expression for the function $F(\gamma)$. It may be verified that this function is represented within 2 percent by

$$F(\gamma)=1-e^{-\gamma}-\gamma e^{-2\gamma}-1.06\gamma^2e^{-3\gamma}+0.28\gamma^2e^{-1.2\gamma}. \quad (2.4)$$

If (2.3) is integrated with the substitution (2.4) we find

$$\frac{w_s}{2ap_0}\left(\frac{\pi}{ct}\right)^{\frac{1}{2}}=\frac{1}{4\pi^{\frac{1}{2}}}\xi\log\left(1+\frac{1}{\xi^2}\right)+\frac{1}{2\pi^{\frac{1}{2}}}\tan^{-1}\xi-\frac{1}{2\pi^{\frac{1}{2}}}\tan^{-1}\xi/2-\frac{1.06\xi}{2\pi^{\frac{1}{2}}(9+\xi^2)}+\frac{0.28\xi}{2\pi^{\frac{1}{2}}(1.44+\xi^2)}. \quad (2.5)$$

By adding a load equal to $p_0/2$ and extending from $x=-\infty$ to $x=+\infty$ we obtain a total uniform load extending from $x=0$ to $x=\infty$ as shown in Fig. 1. In doing so we do not change the settlement because the water is imprisoned in the soil by the impervious layer at the surface and therefore a load extending from $x=-\infty$ to $x=+\infty$ cannot produce any settlement. Hence the settlement for the load p_0 extending over the positive x axis is given by the same expression (2.5) which we write

$$w_s=2ap_0(ct/\pi)^{\frac{1}{2}}f(\xi),$$

where $f(\xi)$ denotes the right-hand side of Eq. (2.5). It is clear that the volume of the water in the soil being constant, the total area under the curve $f(\xi)$ is zero.

In order to represent the settlement as a function of time it is convenient to introduce for the abscissa a characteristic length l which can be chosen arbitrarily. We then have

$$w_s=\frac{2}{\pi^{\frac{1}{2}}}ap_0l\frac{(ct)^{\frac{1}{2}}}{l}f\left[\frac{x}{l}\frac{l}{(ct)^{\frac{1}{2}}}\right]. \quad (2.6)$$

Then settlement curves are plotted as function of x in Fig. 1 at time intervals corresponding to $(ct)^{\frac{1}{2}}/l=1/8,\ 2/8,\ 3/8,\ 4/8,\ 5/8$.

III. Settlement Under a Rectangular Load Distribution

By superposition we may easily derive from the previous solution the settlement due to a constant load p_0 extending from $x=-l/2$ to $x=l/2$. Using (2.6) we get

$$w_s=\frac{2}{\pi^{\frac{1}{2}}}ap_0l\frac{(ct)^{\frac{1}{2}}}{l}\left[f\left(\frac{x+l/2}{(ct)^{\frac{1}{2}}}\right)-f\left(\frac{x-l/2}{(ct)^{\frac{1}{2}}}\right)\right]. \quad (3.1)$$

The settlement curves are represented in Fig. 2 as a function of x at time intervals corresponding to $(ct)^{\frac{1}{2}}/l=1/8,\ 2/8,\ 3/8,\ 4/8,\ 5/8$.

It will be noted that immediately after loading the settlement at the edge is greatly affected by the unloaded section adjacent to it. This unloaded section provides an escape for the water which is not as readily available for the water under the center of the load. Also since this load and deflection are symmetrical to the line $x=0$ no water will flow across this line and therefore the area under the curve from 0 to ∞ must be equal to zero.

If the ratio of the deflection at $x=0$ is compared with the case in a previous paper[2] in which the loading is the same but the water was allowed to flow freely from the top surface the effect of an impervious top surface will be seen. This comparison is made in Fig. 3.

Innovations in Instruments

New Temperature Indicator

Increased precision in measuring thermocouple temperatures, in checking thermocouples, and in measuring small e.m.f.'s, may be obtained without increase in instrument cost by the use of the new portable temperature indicator recently put on the market by the Leeds and Northrup Company, 4934 Stenton Avenue, Philadelphia. In the new indicator 8667, only a small portion of the range is calibrated on a continuously adjustable slidewire; the rest, adjustable in fixed steps, is on a dial switch of ten highly accurate resistors. Range is from 0 to 111 millivolts with a limit of

error of ±0.1 millivolt. The increased accuracy without increase in cost is obtained by sacrificing the convenience of a reference-junction compensator; the only accessories required for this self-contained potentiometer, illustrated in the accompanying figure, are the thermocouple and an ice bath for the reference-junction.

Insulation and Insulating Varnish

The Irvington Varnish and Insulator Company, 24 Argyle Terrace, Irvington, New Jersey, announced the following new products: Irvington Varnished Fiberglas, glass cloth impregnated and coated with heat resisting varnishes, and used as insulation in the manufacture of electric motors, generators and similar equipment; and Harvel 612-C insulating varnish, a phenol-formaldehyde synthetic resin-base varnish made from Cashew Nut Shell Liquid which polymerizes upon baking to an infusible, insoluble state. Literature concerning these products may be obtained by writing the manufacturer.

Variety in Decade Resistance Boxes

Shallcross Manufacturing Company, 10 Jackson Avenue, Collingdale, Pennsylvania, announces a wide range of decade resistance boxes, manufactured in increments from 0.1-ohm steps up to 1,000,000-ohm steps. Some have as many as six decades in a single box. Details are given in Bulletin No. 825.

New Vacuum Tubes

RCA Manufacturing Company, Inc., Harrison, New Jersey, announces that it is making available the following new tubes:

RCA-6SG7 and RCA-12SG7, r-f amplifier pentodes of the metal type particularly recommended for use in high frequency receivers. They feature high transconductance, very low grid-plate capacitance, and two separate cathode terminals. The 6SG7 and 12SG7 are alike except for heater rating. The heater of the 6SG7 is designed so that it can be operated in series with other 6.3-volt, 0.3-ampere types; likewise, the heater of the 12SG7 can be operated in series with other 12.6-volt, 0.15-ampere types.

RCA-45Z3, a miniature, half-wave high vacuum rectifier of the heater-cathode type. Designed specifically for service in a.c./d.c./battery-operated portable receivers, it offers the advantages of small size and low heat dissipation. Its heater requires only 0.075 ampere at 45 volts and its output rating makes it capable of supplying rectified power for both filament and plate circuits in light-weight receivers utilizing the miniature type tubes.

RCA-3Q4, a new, miniature type, power amplifier pentode suitable for operation with 90 volts on plate and screen. It has a power output of 270 milliwatts with 7 percent distortion. The filament of the RCA-3Q4 has a center tap which permits economical usage of this tube in series-filament arrangements with other miniature tubes.

The 6SS7, a remote cut-off, r-f amplifier pentode of the single-ended metal type having a 6.3-volt, 0.15-ampere heater. This new tube provides for a further degree of flexibility in the design of a.c./d.c. receivers utilizing single-ended metal types, where the total heater voltage of a complement of 0.15-ampere types heretofore available would exceed 117 volts.

The 12SN7-GT, a single-ended, twin-triode amplifier having separate cathode terminals for each triode unit. It is recommended for use in resistance-coupled circuits as a voltage amplifier or phase inverter.

RCA-930, new gas photo-tube recommended for use in sound reproduction and relay applications. Electrically, the 930 is like the type 923 with its high sensitivity and large response to red and near infra-red radiation.

The RCA-931, a new type of photo-tube in which the photo-current produced at a light-sensitive cathode is multiplied many times by secondary emission occurring between nine successive dynodes within the tube. It is capable of multiplying feeble currents produced by weak illumination as much as 230,000 times. Focusing of the electron stream is accomplished electrostatically within the tube. The 931 employs the S4 photo-surface which has higher sensitivity to blue-rich light than to blue-deficient light.

The 8001, a transmitting beam pentode suitable for power amplifier, modulator, and oscillator service. The 8001 is particularly useful in all band transmitter designs. In class C telegraph service, the 8001 will provide a power output of approximately 230 watts at frequencies as high as 75 megacycles. Its plate dissipation is 75 watts.

PHOTOGRAPHIC PLATES
for Scientific Purposes

Some sixty kinds of special plates for scientific work are made in Kodak Research Laboratories. Eight basic types of emulsion are used, each carefully prepared to provide desirable combinations of speed, contrast, granularity, and resolving power. They are sensitized in various ways to permit photography in spectral regions ranging from the short-wave-length ultraviolet to the infrared at 12,000 A.

The characteristics of these plates are described in the booklet, *Photographic Plates for Use in Spectroscopy and Astronomy*, 4th Edition; a copy of it will be sent, free, upon request.

EASTMAN KODAK COMPANY

Research Laboratories **Rochester, N. Y.**

A New Text of Unusual Distinction

FUNDAMENTALS OF PHYSICAL SCIENCE

By KONRAD BATES KRAUSKOPF
Assistant Professor of Geology, Stanford University

653 pages, 6 x 9, illustrated. $3.50

In this outstanding new text the sciences of physics, chemistry, astronomy, and geology are presented for the beginning student as a unified field of knowledge, rather than as disconnected summaries of the separate sciences. The ideas and materials in the text have been developed in the well-known pioneering course at Stanford University. The author's unusually broad training in science and his remarkable literary style have enabled him to produce a well balanced and scholarly text that will capture and hold the interest of the beginning student.

Send for a copy on approval

McGRAW-HILL BOOK COMPANY, Inc.

330 West 42nd Street **New York, N. Y.**

New Instrument Booklets

Recent booklets of interest particularly to the physicist include:

The *General Radio Experimenter*, May, 1941, issue, containing articles on "Radio-Frequency Characteristics of the Type 726-A Vacuum-Tube Voltmeter," and "The General Radio Standardizing Laboratory." Published by General Radio Company, 30 State Street, Cambridge, Massachusetts.

Fisher Scientific Company's *The Laboratory*, Vol. 12, No. 4, featuring a description and step-by-step photographs of the procedure for making burettes. Fisher Scientific Company is located at 711–723 Forbes Street, Pittsburgh, Pennsylvania.

Science and Appliance, March, 1941, timely items in science gathered by The Ohio State University Research Foundation, Columbus, Ohio.

The spring issue of *The Educational Focus*, published by Bausch & Lomb Optical Company. The cover of this issue is attractively printed in color.

"Varnished Fiberglas," a folder with samples of black and yellow varnished fiberglas, insulating material manufactured by Irvington Varnish and Insulator Company, 40 Argyle Terrace, Irvington, New Jersey.

The Laboratory, Vol. 12, No. 3, featuring a story on "The Invention of the Mercury Barometer," distributed by Fisher Scientific Company, 711–723 Forbes Street, Pittsburgh, Pennsylvania.

"University Instruments," planned especially for use in the training of engineering students, are described in the leaflet, *Electrical Measurements*, Vol. 6, No. 5, publication of the Sensitive Research Instrument Corporation, 4545 Bronx Boulevard, New York, New York.

Three catalogs recently received from the Roller-Smith Company, 1766 West Market Street, Bethlehem, Pennsylvania, bear the following numbers and titles: 3650, *Outdoor Oil Circuit Breakers;* 3940, *Subway Oil Circuit Breakers;* and 4540, *Precision Balances.*

Brought together for the first time in a single catalog so that it might be more generally available to laboratories, technical schools, universities, and manufacturers, "RCA Measuring Equipment" is described in a 22-page publication recently released by the RCA Manufacturing Company, Inc., Camden, New Jersey. An RCA leaflet has also been recently received discussing "RCA Dynamic Demonstrator, 'The Working Schematic Circuit Diagram.'"

A news-sheet designed to "keep you up to date on relays and relaying methods" is published by the General Electric Company, Schenectady, New York. It is entitled "Relaying News."

Colloidal Graphite as a High Temperature Lubricant and *Colloidal Graphite—Focusing Anode Material for Cathode Ray Tubes*, technical bulletins issued by Acheson Colloids Corporation, Port Huron, Michigan.

Du Mont Oscillographer, February-March, 1941, organ of the Allen B. Du Mont Laboratories, Inc., 2 Main Avenue, Passaic, New Jersey. Contents, "A Single-Sweep Stopwatch and Its Use in a Biological Problem."

"The Successor to Colorimeters. Coleman Universal Spectrophotometer," Bulletin 204 of the Burrell Technical Supply Company, 1936 Fifth Avenue, Pittsburgh, Pennsylvania.

Two new technical bulletins of the Acheson Colloids Corporation, Port Huron, Michigan, entitled: "Colloidal Graphite Lubricants," and "Colloidal Graphite Dispersions."

Wheelco Condensed Catalog, a convenient listing of the principal items of equipment manufactured by the Wheelco Instruments Company, 1929–33 South Halsted Street, Chicago, Illinois.

New laboratory apparatus catalog, XV 41, 174 pages describing instruments and apparatus for science laboratories, available upon application on business or professional stationery, from the American Instrument Company of Silver Spring, Maryland.

"Technical Books on all Subjects, 1941," catalog of the Chemical Publishing Company, Inc., 234 King Street, Brooklyn, New York. Subjects range from "aeronautics" to "foods."

"Vitreosil Pipes and Fittings," Bulletin No. 7, published by The Thermal Syndicate, Ltd., 12 East 46th Street, New York, New York. The vitreous silica pipes and fittings described are "invaluable for extreme chemical, thermal and electrical conditions."

How to apply radiant heat lamps to industrial drying, baking, and heating tasks is told in a new 12-page illustrated booklet A-3817 published by the Westinghouse Lamp Division, Bloomfield, New Jersey. The fundamentals of radiant heat are presented in the first section of this publication, followed by four pages of photographs showing typical drying lamp installations.

The Tuckerman Optical Strain Gauge for measuring tension and compression strains as small as 0.00002 inch in various materials, structural parts and structures, is the subject of the 8-page Bulletin No. 2084 of the American Instrument Company, 8010–8020 Georgia Avenue, Silver Spring, Maryland. The instrument has been on the market for a number of years, but the present bulletin is the most comprehensive yet published.

GAERTNER

INFRA-RED SPECTROGRAPH (Littrow Type)

Range .5461 to 15 microns

with

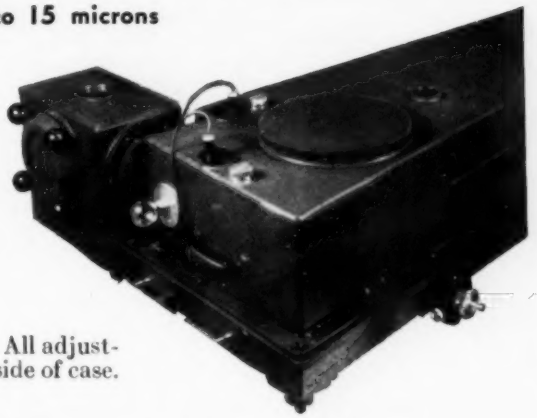
Bilateral Slits of 40 mm clear aperture and provided with micrometer, reading from 0.01 to 3 mm width of opening.

Rocksalt or Potassium Bromide Prisms of best optical quality.

Mirrors are parabolized 10° off axis, focal length 1 m, figured to high accuracy.

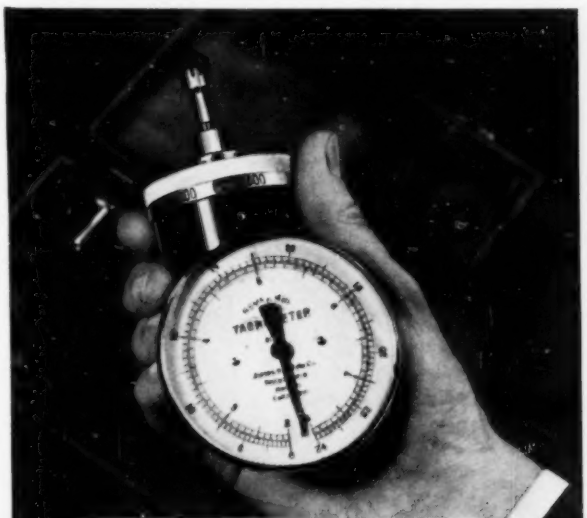
Baffled to eliminate stray light.

Cover is air tight except for the breather tubes. All adjustments of slits, wavelength, etc., are made from outside of case.



THE GAERTNER SCIENTIFIC CORPORATION

1212 Wrightwood Ave. Chicago, U.S.A.



TACHOMETERS

—centrifugal, chronometric and vibrating-reed types. Jagabi Hand Tachoscope and Hand Tachograph. Jagabi Tachonorm.

Write for Bulletin 1505-R

JAMES G. BIDDLE CO.

ELECTRICAL AND SCIENTIFIC INSTRUMENTS

1211-13 ARCH STREET PHILADELPHIA, PA.

PORTABLE POTENTIOMETERS

Rubicon portable potentiometers are available in a variety of single and double range models suitable for thermocouple, pH and instrument calibration work. Guaranteed accuracy to .1%.

These sturdy, self-contained instruments are built to give years of satisfactory service under all conditions encountered in the laboratory, shop or field. They are outstanding values in their class.

Ask for Bulletin 270.

RUBICON COMPANY

Electrical Instrument Makers

Ridge Ave. at 35th Street Philadelphia, Pa.

INDEX TO ADVERTISERS

The following firms have shown their interest in the advancement of physics through their support of the journals and other services of the American Institute of Physics and its Founder Societies. This should entitle them to consideration by physicists whenever possible.

Name *Page*

AMERICAN INSTRUMENT COMPANY viii
Ion Type X-Ray Tube. Powder Spectrum Diffraction Cameras. Scientific Instruments and Supplies.

BAUSCH & LOMB OPTICAL COMPANY .....Cover 3
Makers of Microscopes, Microtomes, Colorimeters, Refractometers, Spectrometers, Balopticons, Photomicrographic and Microprojection Apparatus and related instruments. Also makers of Orthogon Eyeglass Lenses for Better Vision.

JAMES G. BIDDLE COMPANY vii
"Jagabi" Rheostats; Adam Hilger and Kipp & Zonen Optical Instruments; "Pointolite" Lamps; Electrical Testing and Speed-measuring Instruments.

CENTRAL SCIENTIFIC COMPANY .....Cover 4
Manufacturers of Cenco Physical Apparatus and Instruments to meet all requirements of University, College and High School Physics Laboratories. Specializing in high vacuum pumps and development of instruments and apparatus for various sciences.

EASTMAN KODAK COMPANY v
Purified Organic Chemicals for research purposes; Plates for Photography, Photomicrography, Spectroscopy, Photometry, Astronomy; Wratten Light Filters; Cameras and Films.

THE EPPLEY LABORATORY, INC. ii
Standards of e.m.f. (standard cells). Precision electrical instruments; potentiometers, bridges, temperature bridges, volt boxes. Thermopiles and pyrheliometers.

GAERTNER SCIENTIFIC CORPORATION vii
Spectroscopes, Spectrometers, Spectrographs, Spectrophotometers, Heliostats, Measuring Microscopes, Comparators, Cathetometers, Reading Telescopes, Interferometers, Chronographs, Dividing Machines, etc.

GENERAL RADIO COMPANY iii
Manufacturers of electronic measuring instruments; vacuum-tube voltmeters, amplifiers, and oscillators; wave analyzers, noise meters and analyzers, stroboscopes; laboratory standards of capacitance, inductance and frequency; impedance bridges, decade resistors and condensers; air condensers and variable inductors; rheostats, Variacs, transformers; other laboratory accessories.

INDEX—*Continued*

Name *Page*

INTERNATIONAL RESISTANCE COMPANY iv
Manufacturers of metallized and wirewound fixed and variable Resistors including high voltage and high frequency types.

LEEDS & NORTHRUP COMPANY i
Manufacturers of Galvanometers, Resistors, Bridges, Condensers, Inductances, Potentiometers, Testing Sets; Temperature Measuring, Recording and Controlling Apparatus; Instruments for Measuring and Controlling Conductivity of Electrolytes and Hydrogen Ion Concentrations.

MCGRAW-HILL BOOK COMPANY, INC. v

NATIONAL RESEARCH CORPORATION ii
Specialists in industrial applications of high vacuum in the fields of physics and chemistry; research, design, consulting; low reflecting glass surfaces, special products produced under low pressure.

RUBICON COMPANY vii
Galvanometers, electrometers, potentiometers, Wheatstone and Kelvin bridges, resistance boxes, hydrogen ion and conductivity apparatus.

THE SOCIETY OF EXPLORATION GEOPHYSICISTS ii

ABSORPTION CELLS
With Optical Flats Fused On

Windows Flat Over Entire Area to Within 6 Wavelengths

* **Parallelism of liquid-glass interface is ±0.01 mm. in high-precision cells and ±0.025 mm. in standard cells.**
* **Parallelism of the faces of each window is ±10 minutes.**
* **Actual mean length marked on each cell to nearest 0.001 mm.**
* **Actual mean length between faces of windows is ±1% or ±0.005 mm. of the nominal length.**
* **Cells matched to ±0.01 mm. are obtainable.**
* **Pyrex Bodies with Pyrex or Corex "D" Windows. Fused Quartz Bodies and Windows.**

AMINCO

12 Styles
387 Standard Sizes

Fully Described in Bulletin 1170-AP

Strain-free Cells in All Pyrex or All Quartz
Special Cells to Suit Your Requirements

AMERICAN INSTRUMENT CO.
8010 GEORGIA AVENUE · SILVER SPRING, MARYLAND

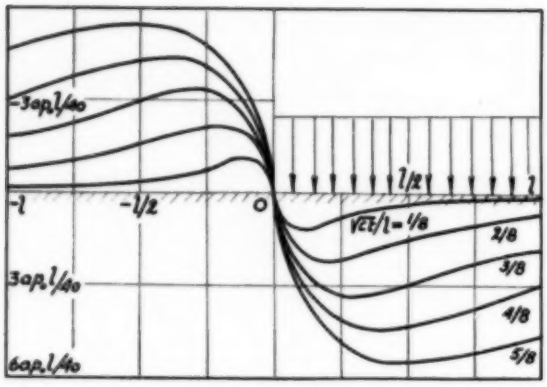


FIG. 1. Settlement of soil surface at various instants of time for a uniform load p_0 extending from $x=0$ to $x=\infty$.

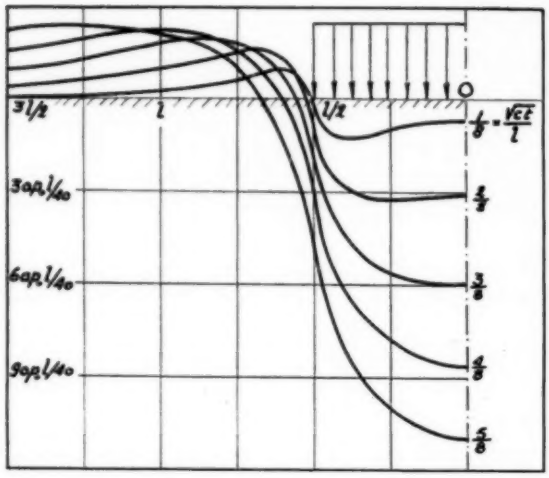


FIG. 2. Settlement of soil surface at various instants of time for a uniform load p_0 extending from $x=-l/2$ to $x=+l/2$.

a tabulated function called the probability integral.

The settlement is composed of two parts: a purely elastic deflection which occurs at the instant of application of the load and a settlement due to consolidation which occurs gradually thereafter. Since we are interested only in the latter we proceed as follows:

For $t=0$ the deflection is purely elastic, its value is

$$w_i = A \sin \lambda x / 2G\lambda. \tag{1.12}$$

The settlement due to consolidation is

$$w_s = w - w_i. \tag{1.13}$$

For convenience we will write Eq. (1.13) in the form

$$w_s = (aA/\lambda) F[\lambda(ct)^{\frac{1}{2}}] \sin \lambda x. \tag{1.14}$$

II. Settlement Under a Uniform Load with Discontinuity

For practical purposes we are interested in the differential settlement occurring when the load distribution has a discontinuity. Such a load distribution will be represented by the following function of x

$$\varphi(x) = \frac{p_0}{\pi} \int_0^\infty \frac{\sin \lambda x}{\lambda} d\lambda \tag{2.1}$$

in which case $\varphi(x) = -p_0/2$ for $x<0$, $p_0/2$ for $x>0$ with a discontinuity at $x=0$.

The settlement due to this load is obtained from (1.14) by the principle of superposition;

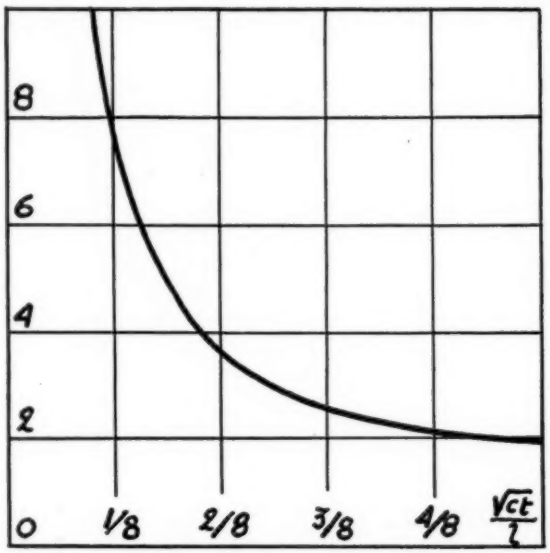


FIG. 3. Rates of the settlement at the center of the loaded area in Fig. 2 when the top surface is pervious to the settlement, when the top surface is impervious.

we find

$$w_s = \frac{ap_0}{\pi} \int_0^\infty \frac{F[\lambda(ct)^{\frac{1}{2}}]}{\lambda^2} \sin \lambda x d\lambda, \tag{2.2}$$

which can be written in the form

$$\frac{w_s}{2ap_0} \left(\frac{\pi}{ct}\right)^{\frac{1}{2}} = \frac{1}{2\pi^{\frac{1}{2}}} \int_0^\infty \frac{F(\gamma)}{\gamma^2} \sin(\gamma\xi) d\gamma \tag{2.3}$$

with $\xi = x/(ct)^{\frac{1}{2}}$.

This integral can be evaluated numerically by the use of an approximate analytical expression for the function $F(\gamma)$. It may be verified that this function is represented within 2 percent by

$$F(\gamma)=1-e^{-\gamma}-\gamma e^{-2\gamma}-1.06\gamma^2e^{-3\gamma}+0.28\gamma^2e^{-1.2\gamma}. \quad (2.4)$$

If (2.3) is integrated with the substitution (2.4) we find

$$\frac{w_s}{2ap_0}\left(\frac{\pi}{ct}\right)^{\frac{1}{2}}=\frac{1}{4\pi^{\frac{1}{2}}}\xi\log\left(1+\frac{1}{\xi^2}\right)+\frac{1}{2\pi^{\frac{1}{2}}}\tan^{-1}\xi-\frac{1}{2\pi^{\frac{1}{2}}}\tan^{-1}\xi/2-\frac{1.06\xi}{2\pi^{\frac{1}{2}}(9+\xi^2)}+\frac{0.28\xi}{2\pi^{\frac{1}{2}}(1.44+\xi^2)}. \quad (2.5)$$

By adding a load equal to $p_0/2$ and extending from $x=-\infty$ to $x=+\infty$ we obtain a total uniform load extending from $x=0$ to $x=\infty$ as shown in Fig. 1. In doing so we do not change the settlement because the water is imprisoned in the soil by the impervious layer at the surface and therefore a load extending from $x=-\infty$ to $x=+\infty$ cannot produce any settlement. Hence the settlement for the load p_0 extending over the positive x axis is given by the same expression (2.5) which we write

$$w_s=2ap_0(ct/\pi)^{\frac{1}{2}}f(\xi),$$

where $f(\xi)$ denotes the right-hand side of Eq. (2.5). It is clear that the volume of the water in the soil being constant, the total area under the curve $f(\xi)$ is zero.

In order to represent the settlement as a function of time it is convenient to introduce for the abscissa a characteristic length l which can be chosen arbitrarily. We then have

$$w_s=\frac{2}{\pi^{\frac{1}{2}}}ap_0l\frac{(ct)^{\frac{1}{2}}}{l}f\left[\frac{x}{l}\frac{l}{(ct)^{\frac{1}{2}}}\right]. \quad (2.6)$$

Then settlement curves are plotted as function of x in Fig. 1 at time intervals corresponding to $(ct)^{\frac{1}{2}}/l=1/8,\ 2/8,\ 3/8,\ 4/8,\ 5/8$.

III. Settlement Under a Rectangular Load Distribution

By superposition we may easily derive from the previous solution the settlement due to a constant load p_0 extending from $x=-l/2$ to $x=l/2$. Using (2.6) we get

$$w_s=\frac{2}{\pi^{\frac{1}{2}}}ap_0l\frac{(ct)^{\frac{1}{2}}}{l}\left[f\left(\frac{x+l/2}{(ct)^{\frac{1}{2}}}\right)-f\left(\frac{x-l/2}{(ct)^{\frac{1}{2}}}\right)\right]. \quad (3.1)$$

The settlement curves are represented in Fig. 2 as a function of x at time intervals corresponding to $(ct)^{\frac{1}{2}}/l=1/8,\ 2/8,\ 3/8,\ 4/8,\ 5/8$.

It will be noted that immediately after loading the settlement at the edge is greatly affected by the unloaded section adjacent to it. This unloaded section provides an escape for the water which is not as readily available for the water under the center of the load. Also since this load and deflection are symmetrical to the line $x=0$ no water will flow across this line and therefore the area under the curve from 0 to ∞ must be equal to zero.

If the ratio of the deflection at $x=0$ is compared with the case in a previous paper[2] in which the loading is the same but the water was allowed to flow freely from the top surface the effect of an impervious top surface will be seen. This comparison is made in Fig. 3.

Innovations in Instruments

New Temperature Indicator

Increased precision in measuring thermocouple temperatures, in checking thermocouples, and in measuring small e.m.f.'s, may be obtained without increase in instrument cost by the use of the new portable temperature indicator recently put on the market by the Leeds and Northrup Company, 4934 Stenton Avenue, Philadelphia. In the new indicator 8667, only a small portion of the range is calibrated on a continuously adjustable slidewire; the rest, adjustable in fixed steps, is on a dial switch of ten highly accurate resistors. Range is from 0 to 111 millivolts with a limit of

error of ±0.1 millivolt. The increased accuracy without increase in cost is obtained by sacrificing the convenience of a reference-junction compensator; the only accessories required for this self-contained potentiometer, illustrated in the accompanying figure, are the thermocouple and an ice bath for the reference-junction.

Insulation and Insulating Varnish

The Irvington Varnish and Insulator Company, 24 Argyle Terrace, Irvington, New Jersey, announced the following new products: Irvington Varnished Fiberglas, glass cloth impregnated and coated with heat resisting varnishes, and used as insulation in the manufacture of electric motors, generators and similar equipment; and Harvel 612-C insulating varnish, a phenol-formaldehyde synthetic resin-base varnish made from Cashew Nut Shell Liquid which polymerizes upon baking to an infusible, insoluble state. Literature concerning these products may be obtained by writing the manufacturer.

Variety in Decade Resistance Boxes

Shallcross Manufacturing Company, 10 Jackson Avenue, Collingdale, Pennsylvania, announces a wide range of decade resistance boxes, manufactured in increments from 0.1-ohm steps up to 1,000,000-ohm steps. Some have as many as six decades in a single box. Details are given in Bulletin No. 825.

New Vacuum Tubes

RCA Manufacturing Company, Inc., Harrison, New Jersey, announces that it is making available the following new tubes:

RCA-6SG7 and RCA-12SG7, r-f amplifier pentodes of the metal type particularly recommended for use in high frequency receivers. They feature high transconductance, very low grid-plate capacitance, and two separate cathode terminals. The 6SG7 and 12SG7 are alike except for heater rating. The heater of the 6SG7 is designed so that it can be operated in series with other 6.3-volt, 0.3-ampere types; likewise, the heater of the 12SG7 can be operated in series with other 12.6-volt, 0.15-ampere types.

RCA-45Z3, a miniature, half-wave high vacuum rectifier of the heater-cathode type. Designed specifically for service in a.c./d.c./battery-operated portable receivers, it offers the advantages of small size and low heat dissipation. Its heater requires only 0.075 ampere at 45 volts and its output rating makes it capable of supplying rectified power for both filament and plate circuits in light-weight receivers utilizing the miniature type tubes.

RCA-3Q4, a new, miniature type, power amplifier pentode suitable for operation with 90 volts on plate and screen. It has a power output of 270 milliwatts with 7 percent distortion. The filament of the RCA-3Q4 has a center tap which permits economical usage of this tube in series-filament arrangements with other miniature tubes.

The 6SS7, a remote cut-off, r-f amplifier pentode of the single-ended metal type having a 6.3-volt, 0.15-ampere heater. This new tube provides for a further degree of flexibility in the design of a.c./d.c. receivers utilizing single-ended metal types, where the total heater voltage of a complement of 0.15-ampere types heretofore available would exceed 117 volts.

The 12SN7-GT, a single-ended, twin-triode amplifier having separate cathode terminals for each triode unit. It is recommended for use in resistance-coupled circuits as a voltage amplifier or phase inverter.

RCA-930, new gas photo-tube recommended for use in sound reproduction and relay applications. Electrically, the 930 is like the type 923 with its high sensitivity and large response to red and near infra-red radiation.

The RCA-931, a new type of photo-tube in which the photo-current produced at a light-sensitive cathode is multiplied many times by secondary emission occurring between nine successive dynodes within the tube. It is capable of multiplying feeble currents produced by weak illumination as much as 230,000 times. Focusing of the electron stream is accomplished electrostatically within the tube. The 931 employs the S4 photo-surface which has higher sensitivity to blue-rich light than to blue-deficient light.

The 8001, a transmitting beam pentode suitable for power amplifier, modulator, and oscillator service. The 8001 is particularly useful in all band transmitter designs. In class C telegraph service, the 8001 will provide a power output of approximately 230 watts at frequencies as high as 75 megacycles. Its plate dissipation is 75 watts.